Equity index variance: evidence from flexible parametric jump–diffusion models

Article (Accepted Version)

Equity Index Variance: Evidence from Flexible Parametric Jump-Diffusion Models

Andreas Kaeck[§] **Paulo Rodrigues**[‡] **Norman J. Seeger**[†]

Abstract. This paper analyzes a wide range of flexible drift and diffusion specifications of stochastic-volatility jump-diffusion models for daily S&P 500 index returns. We find that model performance is driven almost exclusively by the specification of the diffusion component whereas the drift specifications is of second-order importance. Further, the variance dynamics of non-affine models resemble popular non-parametric high-frequency estimates of variance, and their outperformance is mainly accumulated during turbulent market regimes. Finally, we show that jump diffusion models yield more reliable estimates for the expected return of variance swap contracts.

Key Words: Stochastic volatility; jump-diffusion models; Bayesian inference; Markov chain Monte Carlo; particle filter; deviance information criteria; realized variance; high-frequency returns; variance risk premium.

JEL Classifications: C11; G11; G12; G17

[§]University of Sussex, United Kingdom, E-mail: a.kaeck@sussex.ac.uk
[‡]Maastricht University, The Netherlands, E-mail: p.rodrigues@maastrichtuniversity.nl
[†]VU University Amsterdam, The Netherlands, E-mail: n.j.seeger@vu.nl

1 Introduction

This paper analyzes the performance of a wide range of stochastic variance model specifications. Our goal is twofold. First, we aim to study in a very flexible framework the role of various alternative model choices: linear vs non-linear variance drift, linear vs non-linear variance diffusion, Box-Cox transformed variance, and various alternative jump specifications. A large number of models of varying degrees of complexity have been proposed in the literature rather independently of each other.[1] We aim to bring together recent model developments in this strand of literature by analyzing the impact of various extensions on model performance. Our second goal is to provide new out-of-sample evidence. We do this by comparing variance predictions of models estimated on daily return data with non-parametric estimates based on high frequency returns. Recently, new parametric models have been proposed and estimated to high-frequency returns, see Stroud and Johannes (2014) or Bates (2016). To capture stylized facts, such models require several latent state variables and need to account for market microstructure effects such as intradaily seasonality in volatility. Rather than estimating models on high-frequency data, we ask whether models estimated on lower frequency returns are able to generate features calculated from intra-daily returns and which model features are important to do so (see Hansen and Lunde, 2006). Finally, we study the impact of model specification on expected returns of variance swap contracts.

Overall, the literature has reached an understanding that standard affine stochastic volatility models (as, for instance, in Heston, 1993) struggle to explain a range of stylized facts in equity return data, such as sudden sharp price movements or fast-moving variance processes. In order to address some of the shortcomings of standard affine asset price processes, non-

[1] A non-exhaustive list of papers in this strand of the literature includes Bakshi, Cao, and Chen (1997), Andersen, Benzoni, and Lund (2002), Chernov, Gallant, Ghysels, and Tauchen (2003), Eraker, Johannes, and Polson (2003), Jones (2003a,b), Christoffersen, Heston, and Jacobs (2009), Christoffersen, Jacobs, and Mimouni (2010), Egloff, Leippold, and Wu (2010), Bandi and Reno (2016), Bates (2012), Ornthanalai (2014), Yu, Yang, and Zhang (2006), and Zhang and King (2008).

affine drift and diffusion functions of various complexity have been proposed in the literature. Chan, Karolyi, Longstaff, and Sanders (1992), Ait-Sahalia (1996) and Jones (2003a) provide evidence of non-linearities in interest rate dynamics, and Bakshi, Ju, and Ou-Yang (2006), Christoffersen et al. (2010) and Mijatovic and Schneider (2014) among others document support for non-linear model specifications for equity dynamics. Yu et al. (2006) and Zhang and King (2008) follow a different methodology and model Box-Cox transformations of affine variance processes which introduces non-linearities into the return specification via the inverse Box-Cox transformation. One widely used example is the log-variance model, which is a special case of the Box-Cox transformation (see Yu, 2005).

Although the literature has advanced to an understanding that non-affine models may alleviate some of the shortcomings, non-affine models have attracted far less attention in the literature to date and the affine model class of Duffie, Pan, and Singleton (2000) remains very widely used in practice as well as in the academic literature. This is due to the fact that affine models allow for quasi closed-form solutions for European option prices, dynamic asset allocation rules, and transition densities used in econometric estimation. The lack of mathematical tractability has hindered further research on non-affine model dynamics.[2] However, since non-affine models can overcome some of the detriments of affine models it is important to understand the differences between different model classes and how non-affine models improve beyond affine models.

This paper provides several novel contributions. First, we provide a highly flexible modeling approach that encompasses a wide range of specifications previously introduced to the literature and hence we are able to compare the performance of alternative models along various dimensions. Using this general modeling framework we then analyze the extent to which extensions in the drift, diffusion or jump component of the variance process improve model

[2]Note that Aït-Sahalia and Kimmel (2007), Jones (2003b), Benzoni (2002), Christoffersen et al. (2010) and Kaeck and Alexander (2012) analyze option pricing for non-affine stochastic volatility models.

performance. We also estimate and compare Box-Cox transformations of the variance process. Secondly, following Hansen and Lunde (2006), we use non-parametric realized variance (RV) estimated from high-frequency return data as a benchmark for model comparison. This allows us to perform an out-of-sample study of estimated variance paths to test the ability of alternative models to explain the variation in realized variance.[3] RV estimators have also been used as a benchmark in Christoffersen et al. (2010), who use quantile-to-quantile plots to learn about non-affine structures of variance dynamics, and Mijatovic and Schneider (2014) who use RV estimators as a benchmark for variance forecasting performance. However, the focus in our paper is on the comparison of the two variance measures over time. This approach allows us to assess model performance for different market environments (high vs. low volatility) and to visualize the ability of alternative models to cope with extreme market regimes. We provide regression results to study to what extent model-based variances explain the variation in realized variance. Thirdly, we provide a range of new robustness checks. Continuous-time models require computationally time-consuming estimation procedures, which in turn often rule out the possibility of comprehensive robustness checks or rolling window estimations. We provide estimation results for more than 30 models and various subsamples to study model performance over time and to ensure our results are not specific to a particular sample period. We also test the stability of structural model parameters over time to provide evidence of parameter and estimation stability. Fourthly, we investigate the impact of model choice on the estimation of expected returns of popular variance swap contracts. Finally, we investigate how the inclusion of a realized variance measure in the estimation process changes our model ranking results.

Estimation is carried out within a Bayesian statistical framework. We use a Markov Chain Monte Carlo (MCMC) sampling algorithm and apply the Adaptive Rejection Metropolis

[3]This allows us to gage the ability of alternative model specifications to recover model-free variance estimates.

Sampling step (ARMS) proposed by Gilks, Best, and Tan (1995) whenever complete conditional distributions are of unknown form. We use the deviance information criterion (DIC) proposed by Spiegelhalter, Best, Carlin, and van der Linde (2002) to measure model fit based on return data. This measure allows a consistent comparison across different models while taking into account the complexity of the model specification and estimation risk. Our comparison of model-based and realized variances is based on two different variance estimators. As a by-product of model estimation, we first obtain smoothed estimators of all latent state variables (in particular stochastic variance and jump times). However, smoothed variance is estimated by conditioning on the entire return data set and hence is not directly comparable to RV, which is calculated from intra-daily returns only. We therefore also use filtered variance estimators that are consistent with RV in terms of the information set, since they only require index return data up to the time of the estimate.[4]

For the case in which non linearities are captured in the variance process, our most general model specification is closely related to a range of papers on non-affine model dynamics. Our variance models are build on diffusion processes proposed in Ait-Sahalia (1996) who allows for non-linearities in both the drift and diffusion specification to model interest rate dynamics. Bakshi et al. (2006) also build on the framework in Ait-Sahalia (1996) and use variance specifications similar to the ones employed in our work. The empirical analysis in Bakshi et al. (2006), however, differs substantially from ours as their estimation is performed on the VIX index, whereas we estimate all model dynamics from return data. This allows us to obtain consistent model predictions which are unaffected by structural assumptions about the risk premium that is incorporated into VIX dynamics. Following Eraker et al. (2003), our models also allow for jumps in both returns and variance. Our work is also related to a range of papers that extend the dynamics in Heston (1993) to a CEV-type process and use VIX data for model estimation. Jones (2003b) and Aït-Sahalia and Kimmel (2007) analyze

[4]We employ an extension of the particle filter proposed in Johannes, Polson, and Stroud (2009).

a pure stochastic volatility diffusions, Duan and Yeh (2010) also allow for jumps in returns and Kaeck and Alexander (2012) provide further extensions to multi-factor volatility and variance jumps. This literature agrees on the non-linearities in the variance diffusion function, but maintains the assumption of a linear variance drift to retain tractability. Chourdakis and Dotsis (2011) and Mijatovic and Schneider (2014) find evidence for non-linearities in the drift. Similarly, Christoffersen et al. (2010) and Ignatieva, Rodrigues, and Seeger (2015) propose a set of non-affine stochastic volatility specifications and highlight the importance of non-linearity.[5] Our model includes various specifications mentioned above as a special case, by either allowing for more general dynamics of variance or extending the models by the possibility of jumps. In contrast to this literature, Box-Cox transformations for financial time series have been used in Yu et al. (2006) or Zhang and King (2008). To the best of our knowledge our paper is the first that compares both model frameworks.[6]

Our results can be summarized as follows. First, non-affine specifications clearly outperform affine counterparts. Second, we show that non-affine modeling of the variance process outperforms Box-Cox transformed specifications. Third, we show that model performance is almost exclusively driven by the choice of the diffusion specification. The best performing models are equipped with a non-affine diffusion specification and are therefore able to produce large sudden movements in variance, an empirical feature that is also evident in RV paths. This implies that the drift can be modeled with a simple affine function, leading to a model framework with a lower number of parameters. Overall, this also results in a faster and more stable estimation procedure compared to a full general non-affine drift and diffusion model. Furthermore, we observe almost identical model performance during low volatility market regimes, whereas performance differs substantially during times of high market volatility. That

[5] Bandi and Reno (2016) propose a non-parametric model framework that allows for capturing non-affine structures in stochastic volatility. They use daily returns and intraday measures of threshold bipower variation that are based on one-minute price observations to estimate the variance.

[6] We thank an anonymous referee for the suggestion to add this model class as an alternative way to model non-linearities in stochastic processes for stock returns.

is, model complexity is most beneficial during periods of market turmoil. Moreover, we show that incorporating jumps into the models improves the estimation of expected variance swap returns. Consistent with our earlier findings, the diffusion part of the specification his the main driver of these results. Finally, we show that our conclusion remain valid if we extend the information set in the estimation and estimate the models using both daily returns as well as a realized variance measure.

2 Model Description

For our benchmark models, we assume that the logarithm of the index value $Y_t = \ln(S_t)$ and its diffusive variance V_t solve the following system of stochastic differential equations:

$$dY_t = \mu\, dt + \sqrt{V}_t\, dW_t^y + d\left(\sum_{j=1}^{N_t} \xi_j^y\right) \tag{1}$$

$$dV_t = \alpha(V_t)\, dt + \beta(V_t)\, dW_t^v + d\left(\sum_{j=1}^{N_t} \xi_j^v\right) \tag{2}$$

where dW_t^y and dW_t^v denote Brownian motion increments with non-zero correlation, i.e. $E(dW_t^y dW_t^v) = \rho\, dt$.[7] We assume a constant drift term μ in the log process, and N_t denotes a Poisson process with constant jump intensity λ. Jumps in the state variables are contemporaneous, as we assume that N_t enters both the return and variance equation (hence we follow the literature in assuming simultaneous jumps, see Eraker, 2004). Jump sizes in variance ξ^v follow an exponential distribution with expectation μ_V, i.e. $\xi_t^v \sim \mathcal{E}\left(\mu_V\right)$, and jumps in returns are conditionally normally distributed with mean $\mu_Y + \rho_J\, \xi_t^v$ and variance σ_Y^2, i.e. $\xi_t^y \,|\, \xi_t^v \sim \mathcal{N}\left(\mu_Y + \rho_J \xi_t^v, \sigma_Y^2\right)$. Finally, the functions $\alpha(V_t)$ and $\beta(V_t)$ specify, respectively, the

[7] We use the standard notation $E(\cdot)$ for the expectation.

drift and the diffusion of the variance. In the most general form, these functions are given by

$$\alpha(V_t) = \alpha_0 + \alpha_1 \frac{1}{V_t} + \alpha_2 V_t + \alpha_3 V_t^2, \tag{3}$$

$$\beta(V_t) = \beta_0 + \beta_1 V_t + \beta_2 V_t^{\beta_3}. \tag{4}$$

This general specification nests many continuous-time models used in the literature. Restricting the jump intensity to zero results in a stochastic volatility (SV) model with continuous sample paths. In this model class, the unrestricted drift and diffusion specification for the variance (called *PolyPoly*) is used in Ait-Sahalia (1996) to analyze short rate models.[8] Conley, Hansen, Luttmer, and Scheinkman (1997) and Chourdakis and Dotsis (2011) use a specification where the diffusion parameters β_0 and β_1 are set to zero (labeled *PolyCev*) when analyzing short rate models and equity returns. Further restricting the drift parameters α_1 and α_3 to zero (called *AffineCev*) results in the SV model used in Jones (2003b) for the analysis of equity index dynamics. Fixing the parameter β_3 to either $\frac{1}{2}$, 1 or $\frac{3}{2}$ results in the models analyzed in Christoffersen et al. (2010). These diffusion assumptions are labeled *Sqr*, *One* and *3/2*. Finally, restricting the parameters α_1, α_3, β_0 and β_1 to zero, and fixing the parameter β_3 at $\frac{1}{2}$, results in the stochastic volatility model analyzed in Heston (1993). This specification is the only fully affine model, which we label *AffineSqr*. We refer to these model specifications as 'variance models'.

Alternatively, we use a Box-Cox transformation to model non-linearities and assume that the logarithm of the index value $Y_t = \ln(S_t)$ and the Box-Cox transformation of the process

[8] In their most general form the functions α and β are not polynomial, but we nevertheless follow Ignatieva et al. (2015) in using this name convention.

H_t solve the following system of stochastic differential equations

$$dY_t = \mu \, dt + \sqrt{g(H_t)} \, dW_t^y + d\left(\sum_{j=1}^{N_t} \xi_j^y\right) \tag{5}$$

$$dH_t = (\alpha_0 + \alpha_2 H_t) \, dt + \beta_0 \, dW_t^h + d\left(\sum_{j=1}^{N_t} \xi_j^v\right) \tag{6}$$

where

$$H_t = \begin{cases} \left(V_t^\delta - 1\right) / \delta & \text{if } \delta \neq 0 \\ \ln(V_t) & \text{if } \delta = 0 \end{cases}$$

and the inverse transformation being given by

$$g(H_t) = \begin{cases} (1 + \delta H_t)^{1/\delta} & \text{if } \delta \neq 0 \\ \exp(H_t) & \text{if } \delta = 0. \end{cases}$$

Note that the variable H_t is modeled as an affine process but the overall specification is non-affine due to the non-linear transformation $g(H_t)$ in the return process. The additional parameter δ controls the degree of non-linearity in the return process. As shown in Yu et al. (2006) the model reduces to several known stochastic variance specifications for different parameter combinations. We have also experimented with further non-affine extensions of the process H_t, but find that such models are over-parametrized.[9] We refer to these model specifications as 'Box-Cox variance models'.

Including a jump component in the return process allows us to model large return outliers, and the model class with stochastic volatility and jumps in prices (labeled SVJ) has attracted considerable interest in the literature. Bates (1996) for instance uses an affine version of this

[9] One special case of the Box-Cox transformation is the log variance processes, which is given by restricting the Box-Cox parameter to $\delta = 0$.

model to analyze the pricing of exchange rate options. Including a jump component in the return and the variance process leads to the SVCJ model introduced in Duffie et al. (2000), a specification further studied in Eraker et al. (2003), Eraker (2004) and others. We employ a constant jump intensity framework in our analysis, a set-up frequently used in the literature (see Eraker et al., 2003, Broadie, Chernov, and Johannes, 2007, Ferriani and Pastorello, 2012 and Durham, 2013). We note that the assumption of a constant jump intensity λ can be relaxed, as it has been in a number of studies, allowing for more flexibility but at the same time introducing more complexity. For example, Bates (2000) and Pan (2002) assume that the jump intensity is an affine function of return variance, Santa-Clara and Yan (2010) estimate a model in which the jump intensity follows a stochastic process that is independent of the variance, and Durham and Park (2013) assume a Markov switching model for the intensity. Since the focal point of this paper is to study the importance of the drift vs the diffusion function of the variance process, we avoid additional complexity by assuming constant jump intensities. Equipped with three model classes (SV, SVJ, SVCJ), two variance drift specifications (*Poly*, *Affine*), five variance diffusion specifications (*Cev*, *Sqr*, *One*, *3/2*, *Poly*), and the Box-Cox transformation we analyze a total of thirty-three models which are listed in Table 1. For ease of exposition, we only present results for a subset of models and remove detailed results whenever the model performance is indistinguishable from related specifications. The complete set of results is available from the authors on request.

3 Estimation Methodology

3.1 Discretization

For model estimation, we follow Eraker et al. (2003) and use a standard Euler scheme with discretization interval Δ set to one trading day, i.e. $\Delta = 1$. Denoting $R_t = 100 \times (Y_t - Y_{t-1})$

Table 1: Overview of Models

This table shows the different specifications of the drift and diffusion terms for the dynamics of the stochastic variance. For each model class SV, SVJ, and SVCJ, we estimate all specifications provided in this table.

Drift	Diffusion	Features
Affine	Sqr	variance drift is affine in variance, square root diffusion ($\beta_3 = 0.5$)
Affine	One	variance drift is affine in variance, linear diffusion ($\beta_3 = 1$)
Affine	3/2	variance drift is affine in variance, 3/2 diffusion ($\beta_3 = 1.5$)
Affine	Cev	variance drift is affine in variance, free diffusion ($\beta_3 \in [0.5; 1.5]$)
Affine	Sqr	variance drift is affine in variance, diffusion is polynomial
Poly	Sqr	variance drift is polynomial in variance, square root diffusion ($\beta_3 = 0.5$)
Poly	One	variance drift is polynomial in variance, linear diffusion ($\beta_3 = 1$)
Poly	3/2	variance drift is polynomial in variance, 3/2 diffusion ($\beta_3 = 1.5$)
Poly	Cev	variance drift is polynomial in variance, free diffusion ($\beta_3 \in [0.5; 1.5]$)
Poly	Poly	variance drift is polynomial in variance, diffusion is polynomial
BoxCox		Box-Cox transformed process is affine, transformation into return is non-linear

as the daily percentage log return of the S&P 500 index, the discretized version of the system of equations (1) and (2) can be expressed as

$$R_t = \mu + \sqrt{V_{t-1}}\,\varepsilon_t^y + \xi_t^y J_t \tag{7}$$

$$V_t = V_{t-1} + \alpha_0 + \alpha_1 \frac{1}{V_{t-1}} + \alpha_2 V_{t-1} + \alpha_3 V_{t-1}^2 + \left(\beta_0 + \beta_1 V_{t-1} + \beta_2 V_{t-1}^{\beta_3}\right)\varepsilon_t^v + \xi_t^v J_t\,, \tag{8}$$

where shocks to the return and variance equation are given by $\varepsilon_t^y = W_t^y - W_{t-1}^y$ and $\varepsilon_t^v = W_t^v - W_{t-1}^v$, and follow a bivariate normal distribution with zero expectation, unit variance and correlation ρ. In the Euler discretization scheme, we limit the number of jumps per day to a maximum of one, hence we set the indicator J_t to one in the event of a jump (which occurs with probability λ) and equal to zero in the case of no jump. Note that the jump indicator J_t in the return equation is identical to the indicator in the variance equation, since jumps occur simultaneously. The jump sizes retain the distributional assumptions described in Section 2.[10] For technical details regarding the discretization schemes and the existence of stationary distributions of the models, as well as simulation results, the reader is referred to Ait-Sahalia (1996), Conley et al. (1997), Eraker et al. (2003), Jones (2003a) and Jones (2003b). The discretized version of the system of equations (5) and (6) can be expressed as

$$
\begin{aligned}
R_t &= \mu + \sqrt{g(H_{t-1})}\,\varepsilon_t^y + \xi_t^y J_t && (9)\\
H_t &= H_{t-1} + \alpha_0 + \alpha_2 H_{t-1} + \beta_0 \varepsilon_t^h + \xi_t^v J_t\,, && (10)
\end{aligned}
$$

where shocks to the Box-Cox transformed variance equation are given by $\varepsilon_t^h = W_t^h - W_{t-1}^h$.

The simple Euler scheme may result in a discretization bias in our estimation. We know from prior literature, i.e., Eraker et al. (2003), Li, Wells, and Yu (2008), and Yu et al. (2006) that this bias is negligible for affine model specifications and the Box-Cox transformed variance models. It is therefore unlikely that the Euler scheme produces a substantial bias when estimating non-affine jump-diffusion models. In order to rule out this possibility we extend the simulation results in the literature. To this end, we choose 'true' model parameters similar to estimates reported in the related literature. Using an Euler discretization with 100 time steps per day, we first simulate 100 artificial sample paths which are used to calculate 4000 daily return observations. Our results in Table 15 in the Appendix show that for all three tested models (*SV PolyPoly*, *SVJ PolyPoly* and *SVCJ PolyPoly*) the estimation methodology is very accurate and the time-discretization does not introduce any notable bias in the estimation results. All root mean squared errors (RMSE) are very low and even the estimation of jump parameters that are notoriously difficult to estimate due to their rare-event character provide satisfactory results.

[10] The assumption of at most one jump per day could lead to some discretization bias when estimating jump parameters. However, the following example demonstrates that since jumps are rare events, discretization bias is typically very small. Using $P(N_t - N_{t-1} = j) = \frac{\exp\{-\lambda\}\lambda^j}{j!}$ and assuming the jump intensity to be $\lambda = 0.1$, the probability of observing more than one jump per day is 0.0047. Note that our estimation results indicate estimates for λ much smaller than 0.1.

3.2 Estimation

Because of their advantages for estimating models with latent state variables, we employ Bayesian estimation and model testing methods based on Markov Chain Monte Carlo (MCMC) sampling algorithms. In the context of estimating equity return models, MCMC methods were pioneered by Jacquier, Polson, and Rossi (1994) and Jacquier, Polson, and Rossi (2004) and have subsequently been successfully applied to a plethora of models as well as to different financial and economic time-series data. Below, we provide a brief overview of the sampling algorithm for the *PolyPoly* model in the SVCJ class, since this specification is the most complex in our analysis. Estimation algorithms of the nested models follow accordingly. For a general introduction to MCMC methods, the reader is referred to Casella and George (1992), Chib and Greenberg (1995), and Johannes and Polson (2009).

Bayes' theorem implies that the posterior distribution of the parameters and the latent states is proportional to the likelihood times the prior distribution. Using the Euler discretization, it follows that the posterior distribution is proportional to

$$\prod_{t=1}^{T} p(R_t, V_t|V_{t-1}, \xi_t^y, \xi_t^v, J_t, \mathbf{\Theta}) \times p(\xi_t^y|\xi_t^v, \mu_Y, \rho_J, \sigma_Y) \times p(\xi_t^v|\mu_V) \times p(J_t|\lambda) \times p(\mathbf{\Theta}) \tag{11}$$

where the vector $\mathbf{\Theta} = \{\mu, \rho, \alpha_0, \alpha_1, \alpha_2, \alpha_3, \beta_0, \beta_1, \beta_2, \beta_3, \lambda, \mu_J, \sigma_J, \rho_J, \mu_V\}$ collects all model parameters. We assume independent parameter priors so that the prior for the full parameter vector $p(\mathbf{\Theta})$ can be decomposed into the product of univariate parameter priors. We follow Eraker et al. (2003) and use priors with very large variances such that our results are solely driven by the information in the data and not by assumptions about the prior distribution. We refer the reader to Eraker et al. (2003) regarding details on the assumed prior distributions.

We decompose the high-dimensional posterior into one-dimensional complete conditional distributions and draw every parameter and latent variable individually. We use conjugate prior distributions whenever possible; in particular, we draw the parameter μ, all α parameters, and latent states and parameters of the jump processes from known distributions. For the derivation of the complete conditionals in these cases, the reader is referred to Ignatieva et al. (2015). To draw parameters with unknown complete conditional distribution (β's, ρ, and the variances) we use the Adaptive Rejection Metropolis Sampling (ARMS) algorithm developed in Gilks et al. (1995). This algorithm is computationally intensive but provides very advantageous estimation results, as reported in Li et al. (2008), as rejection rates in the MCMC algorithm are typically extremely low and the mixing of the chain is comparable to the case when parameters can be directly drawn from a complete conditional distribution.[11]

For the Box-Cox transformed model specifications we employ the same Bayesian estimation methodology. The structure of the model, however, leads to further complexity in the estimation algorithm to ensure that the daily variance implied by the draw of the Box-Cox transformation exists. Note that

$$g(H_t) = \begin{cases} (1+\delta H_t)^{1/\delta} & \text{if } \delta \neq 0 \\ \exp(H_t) & \text{if } \delta = 0. \end{cases}$$

In order to ensure positivity of the variance process when drawing the Box-Cox transformed variables H_t, interrelations between δ and H_t have to be taken into account. In particular, since δ is negative in our estimation we have to make sure that $(1+\delta H_t) > 0$ for V_t to be real-valued. Furthermore, the draw of δ given the time series of H_t also needs to ensure the existence of V_t, i.e., $(1+\delta H_t) > 0$. We account for this by truncating the proposal for H_t in our draws to be of the set $H_t > -1/\delta$ and restrict proposals for δ to neighborhood of previous draws.

[11]Rejection rates are indeed equal to zero for log-concave densities.

3.3 Model Comparison

We compare the models under consideration along two dimensions, namely model fit and estimated variance paths. As a measure of model fit, we use the deviance information criterion (DIC) derived in Spiegelhalter et al. (2002). DIC is employed to compare stochastic variance models for equity index returns by Berg, Meyer, and Yu (2004) who show in a simulation study that DIC is capable of adequately ranking competing stochastic volatility models for equity returns. Similar to other non-Bayesian fit statistics, DIC penalizes model complexity and rewards model fit. DIC is defined as

$$\begin{aligned} DIC &= \bar{D} + p_D \\ &= 2E_{\Theta|R}[-2\ln(p(R|\mathbf{\Theta}))] + 2\ln(p(R|\bar{\mathbf{\Theta}})). \end{aligned} \tag{12}$$

with the components $\bar{D}$ and p_D defined as

$$\begin{aligned} \bar{D} &= E_{\Theta|R}[-2\ln(p(R|\mathbf{\Theta}))] \\ p_D &= E_{\Theta|R}[-2\ln(p(R|\mathbf{\Theta}))] + 2\ln(p(R|\bar{\mathbf{\Theta}})), \end{aligned}$$

where $\ln(p(R|\bar{\mathbf{\Theta}}))$ denotes the log-likelihood function evaluated at the posterior mean and $E_{\Theta|R}(\ln(p(R|\mathbf{\Theta})))$ denotes the posterior mean of the log-likelihood function.[12] The computation of the statistic is readily obtained via the MCMC estimation output. The model with the lowest DIC can be interpreted as the model with the best return prediction.

Additionally, we analyze the variance dynamics for each model, and compare estimated variance paths to a high-frequency realized variance (RV) estimator of daily quadratic variation as proposed in Shephard and Sheppard (2010), Andersen, Bollerslev, Diebold, and Ebens (2001a), and Andersen, Bollerslev, Diebold, and Labys (2001b). We download the estimator

[12]For details of computation we refer to Spiegelhalter et al. (2002) and Berg et al. (2004)

from the Oxford Man institute and use the series "5 minute returns with 1 minute subsampling" to account for microstructure noise. As in Andersen, Fusari, and Todorov (2015) and Hansen and Lunde (2006) we add the quadratic overnight return to arrive at an estimator for the close to close quadratic variation. This estimator provides a non-parametric benchmark to which we compare our daily variance paths from parametric models. A potential inconsistency arises from the fact that high-frequency estimators only use information up to time t whereas posterior means from the MCMC algorithm are based on information up to time $T \geq t$. A filtering method is therefore employed to extract a non-forward-looking estimator of model-implied variance based on return data up to time t only. We use the filtering algorithm proposed by Johannes et al. (2009) which adapts the auxiliary particle filter to continuous-time models, and we adjust it to our non-affine model framework. In general, the algorithm consists of three steps.[13] First, particles generated in $t-1$ are resampled. Secondly, the latent variables are propagated forward using the resampled latent states.[14] And thirdly, the resulting particles are re-weighted using an importance-sampling scheme. This step is needed due to the use of approximated distributions in the first two steps.

3.4 Model Implementation

The MCMC algorithm is implemented in C++ using random number generators of the GNU Scientific Library. The Markovian dependence structure of the variances can be used to draw variances in blocks of two (as described in Jones, 2003b), and this offers the possibility of some performance gain by parallelization. The mixing of the chain depends heavily on the model specification. For the SV-AffineSqr model, convergence is obtained relatively quickly after 30,000 draws following a burn-in period of 20,000 draws, whereas more complex models,

[13]For full details on the algorithm we refer to Johannes et al. (2009)

[14]To propagate the variances forward we use the Euler discretization given in Equation (8) and use the results in Johannes et al. (2009) to draw the jump times and jump sizes.

such as SVCJ-PolyCev, require more simulation runs. After extensive testing of parameter convergence we base our estimation results for all models on 150,000 draws following a burn-in period of 50,000 draws. Is it particularly interesting to note that calculating DIC statistics, and therefore a stable model ranking, is much more sensitive to the number of MCMC draws than parameter estimation. To receive a stable DIC ranking we run our algorithm with 9,000,000 draws following a burn-in period of 900,000 draws. Hence, all reported model ranking results are based on DIC statistics that are based on 9,000,000 draws whereas model parameters utilize 150,000 draws. The run time for a full model estimation is strongly dependent on the model specification. All calculations are performed on a large computer cluster equipped with Intel Xeon L5520 2.26 GHz processors.

4 Empirical Results

4.1 Model Estimation

We first estimate all model parameters and use daily percentage returns of the S&P 500 index obtained from CRSP from January 1983 until December 2013. Tables 2, 3 and 4 summarize the estimation results for SV, SVJ and SVCJ models over the whole sample period. For affine and simple non-affine extensions, these parameters have been discussed extensively in the literature and our parameter estimates for these models confirm previous findings. We therefore confine our discussion to a few noteworthy results.

Table 2: Parameter Estimators for the SV Model Class

This table shows posterior means and standard deviations (in brackets) of model parameters for all drift and diffusion specifications of the SV class. Parameter estimation is based on daily S&P 500 percentage returns from January 1983 until December 2013.

Par	**AffineSqr**	**PolySqr**	**AffineOne**	**PolyOne**	**Affine3/2**	**Poly3/2**	**AffineCev**	**PolyCev**	**AffinePoly**	**PolyPoly**	**BoxCox**
μ	0.0307	0.0315	0.0371	0.0371	0.0506	0.0515	0.0382	0.0384	0.0363	0.0382	0.0392
	(0.0089)	(0.0088)	(0.0089)	(0.0088)	(0.0086)	(0.0087)	(0.0089)	(0.0089)	(0.0089)	(0.0089)	(0.0089)
α_0	0.0246	0.0020	0.0165	0.0122	0.0034	0.0031	0.0128	0.0109	0.0148	0.0088	-0.0081
	(0.0028)	(0.0019)	(0.0023)	(0.0032)	(0.0006)	(0.0007)	(0.0032)	(0.0029)	(0.0031)	(0.0033)	(0.0022)
α_1	–	0.0059	–	0.0007	–	0.0001	–	0.0005	–	0.0013	–
	–	(0.0008)	–	(0.0006)	–	(0.0001)	–	(0.0005)	–	(0.0008)	–
α_2	-0.0212	-0.0110	-0.0117	-0.0060	-0.0006	-0.0006	-0.0067	-0.0042	-0.0097	-0.0038	-0.0203
	(0.0027)	(0.0024)	(0.0035)	(0.0037)	(0.0006)	(0.0005)	(0.0045)	(0.0034)	(0.0045)	(0.0032)	(0.0030)
α_3	–	-0.0003	–	-0.0014	–	-0.0004	–	-0.0014	–	-0.0013	–
	–	(0.0002)	–	(0.0008)	–	(0.0003)	–	(0.0008)	–	(0.0008)	–
β_0	–	–	–	–	–	–	–	–	0.0180	0.0015	0.1759
	–	–	–	–	–	–	–	–	(0.0065)	(0.0014)	(0.0119)
β_1	–	–	–	–	–	–	–	–	0.0154	0.1455	–
	–	–	–	–	–	–	–	–	(0.0094)	(0.0425)	–
β_2	0.1685	0.1892	0.1933	0.1942	0.1236	0.1237	0.1846	0.1883	0.1438	0.0415	–
	(0.0085)	(0.0089)	(0.0115)	(0.0115)	(0.0069)	(0.0069)	(0.0127)	(0.0125)	(0.0167)	(0.0422)	–
β_3	–	–	–	–	–	–	1.0905	1.0642	1.2126	1.3772	–
	–	–	–	–	–	–	(0.0668)	(0.0597)	(0.0756)	(0.3846)	–
ρ	-0.5997	-0.6153	-0.6143	-0.6150	-0.5909	-0.5870	-0.6145	-0.6141	-0.6177	-0.6147	-0.5868
	(0.0334)	(0.0329)	(0.0334)	(0.0333)	(0.0392)	(0.0391)	(0.0343)	(0.0339)	(0.0338)	(0.0337)	(0.0364)
δ	–	–	–	–	–	–	–	–	–	–	-0.2435
	–	–	–	–	–	–	–	–	–	–	(0.0528)

Table 3: Parameter Estimators for the SVJ Model Class

This table shows posterior means and standard deviations (in brackets) of model parameters for all drift and diffusion specifications of the SVJ class. Parameter estimation is based on daily S&P 500 percentage returns from January 1983 until December 2013.

Par	**AffineSqr**	**PolySqr**	**AffineOne**	**PolyOne**	**Affine3/2**	**Poly3/2**	**AffineCev**	**PolyCev**	**AffinePoly**	**PolyPoly**	**BoxCox**
μ	0.0325	0.0324	0.0426	0.0430	0.0625	0.0633	0.0470	0.0466	0.0436	0.0457	0.0473
	(0.0089)	(0.0087)	(0.0090)	(0.0090)	(0.0090)	(0.0090)	(0.0093)	(0.0092)	(0.0092)	(0.0091)	(0.0093)
α_0	0.0198	0.0017	0.0131	0.0095	0.0024	0.0022	0.0084	0.0079	0.0101	0.0065	-0.0071
	(0.0025)	(0.0016)	(0.0020)	(0.0027)	(0.0005)	(0.0006)	(0.0022)	(0.0022)	(0.0024)	(0.0025)	(0.0020)
α_1	–	0.0051	–	0.0007	–	0.0001	–	0.0004	–	0.0010	–
	–	(0.0007)	–	(0.0005)	–	(0.0001)	–	(0.0004)	–	(0.0006)	–
α_2	-0.0171	-0.0089	-0.0087	-0.0042	-0.0005	-0.0005	-0.0031	-0.0024	-0.0051	-0.0024	-0.0152
	(0.0024)	(0.0022)	(0.0031)	(0.0029)	(0.0005)	(0.0005)	(0.0027)	(0.0022)	(0.0034)	(0.0021)	(0.0024)
α_3	–	-0.0003	–	-0.0012	–	-0.0003	–	-0.0011	–	-0.0010	–
	–	(0.0002)	–	(0.0007)	–	(0.0003)	–	(0.0007)	–	(0.0007)	–
β_0	–	–	–	–	–	–	–	–	0.0149	0.0011	0.1592
	–	–	–	–	–	–	–	–	(0.0060)	(0.0011)	(0.0107)
β_1	–	–	–	–	–	–	–	–	0.0104	0.1317	–
	–	–	–	–	–	–	–	–	(0.0065)	(0.0372)	–
β_2	0.1505	0.1742	0.1778	0.1794	0.1148	0.1148	0.1653	0.1707	0.1338	0.0379	–
	(0.0085)	(0.0091)	(0.0112)	(0.0113)	(0.0065)	(0.0066)	(0.0120)	(0.0120)	(0.0146)	(0.0363)	–
β_3	–	–	–	–	–	–	1.1467	1.1007	1.2556	1.4963	–
	–	–	–	–	–	–	(0.0615)	(0.0606)	(0.0708)	(0.3400)	–
ρ	-0.6590	-0.6751	-0.6754	-0.6758	-0.6527	-0.6485	-0.6768	-0.6742	-0.6788	-0.6744	-0.6743
	(0.0310)	(0.0302)	(0.0310)	(0.0310)	(0.0371)	(0.0372)	(0.0322)	(0.0316)	(0.0316)	(0.0316)	(0.0339)
λ	0.0044	0.0046	0.0110	0.0113	0.0181	0.0175	0.0135	0.0126	0.0123	0.0123	0.0172
	(0.0016)	(0.0018)	(0.0052)	(0.0055)	(0.0076)	(0.0074)	(0.0064)	(0.0059)	(0.0057)	(0.0057)	(0.0077)
μ_Y	-2.8786	-2.7379	-1.3741	-1.3259	-0.9077	-0.9347	-1.1794	-1.2168	-1.2385	-1.2190	-1.0212
	(0.8602)	(0.8679)	(0.5432)	(0.5261)	(0.3780)	(0.3931)	(0.4861)	(0.4891)	(0.4872)	(0.4842)	(0.4062)
σ_Y	2.9857	2.8204	1.7436	1.7290	1.6208	1.6326	1.6556	1.6754	1.6760	1.6912	1.5841
	(0.5929)	(0.5572)	(0.2824)	(0.2700)	(0.2064)	(0.2098)	(0.2335)	(0.2385)	(0.2363)	(0.2421)	(0.2084)
δ	–	–	–	–	–	–	–	–	–	–	-0.2991
	–	–	–	–	–	–	–	–	–	–	(0.0611)

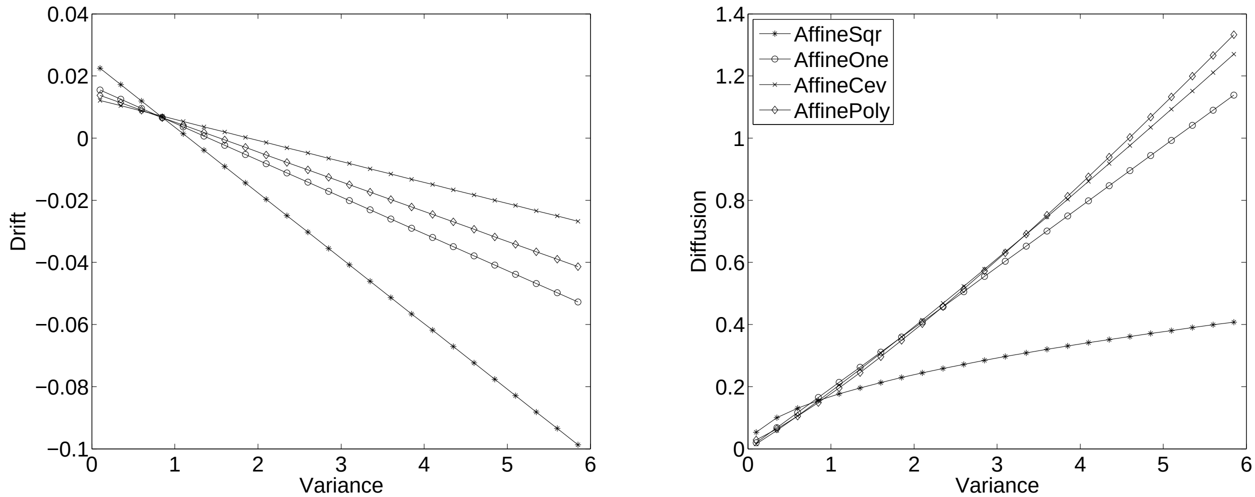
Table 4: Parameter Estimators for the SVCJ Model Class

This table shows posterior means and standard deviations (in brackets) of model parameters for all drift and diffusion specifications of the SVCJ class. Parameter estimation is based on daily S&P 500 percentage returns from January 1983 to December 2013.

Par	**AffineSqr**	**PolySqr**	**AffineOne**	**PolyOne**	**Affine1.5**	**Poly1.5**	**AffineCev**	**PolyCev**	**AffinePoly**	**PolyPoly**	**BoxCox**
μ	0.0387	0.0364	0.0421	0.0421	0.0533	0.0527	0.0466	0.0452	0.0432	0.0427	0.0467
	(0.0089)	(0.0087)	(0.0089)	(0.0088)	(0.0088)	(0.0089)	(0.0092)	(0.0092)	(0.0092)	(0.0090)	(0.0091)
α_0	0.0183	0.0021	0.0109	0.0060	0.0015	0.0007	0.0051	0.0034	0.0090	0.0035	-0.0110
	(0.0025)	(0.0019)	(0.0019)	(0.0031)	(0.0008)	(0.0006)	(0.0028)	(0.0028)	(0.0034)	(0.0026)	(0.0024)
α_1	–	0.0043	–	0.0009	–	0.0002	–	0.0007	–	0.0014	–
	–	(0.0008)	–	(0.0006)	–	(0.0001)	–	(0.0005)	–	(0.0007)	–
α_2	-0.0241	-0.0133	-0.0124	-0.0070	-0.0033	-0.0041	-0.0078	-0.0057	-0.0108	-0.0059	-0.0153
	(0.0031)	(0.0026)	(0.0031)	(0.0037)	(0.0024)	(0.0023)	(0.0038)	(0.0035)	(0.0042)	(0.0036)	(0.0025)
α_3	–	-0.0003	–	-0.0010	–	-0.0004	–	-0.0008	–	-0.0009	–
	–	(0.0003)	–	(0.0007)	–	(0.0004)	–	(0.0006)	–	(0.0006)	–
β_0	–	–	–	–	–	–	–	–	0.0096	0.0015	0.1410
	–	–	–	–	–	–	–	–	(0.0062)	(0.0013)	(0.0114)
β_1	–	–	–	–	–	–	–	–	0.0092	0.0995	–
	–	–	–	–	–	–	–	–	(0.0066)	(0.0442)	–
β_2	0.1275	0.1559	0.1584	0.1600	0.1203	0.1091	0.1415	0.1483	0.1290	0.0560	–
	(0.0090)	(0.0104)	(0.0114)	(0.0114)	(0.0123)	(0.0079)	(0.0135)	(0.0141)	(0.0147)	(0.0423)	–
β_3	–	–	–	–	–	–	1.2333	1.1567	1.1998	1.1792	–
	–	–	–	–	–	–	(0.1091)	(0.1254)	(0.1183)	(0.3515)	–
ρ	-0.6743	-0.6921	-0.7169	-0.7200	-0.7806	-0.8416	-0.7982	-0.7771	-0.7391	-0.7373	-0.7028
	(0.0347)	(0.0330)	(0.0376)	(0.0389)	(0.0711)	(0.0422)	(0.0597)	(0.0625)	(0.0530)	(0.0491)	(0.0384)
λ	0.0056	0.0064	0.0087	0.0091	0.0161	0.0165	0.0157	0.0144	0.0106	0.0109	0.0099
	(0.0018)	(0.0021)	(0.0034)	(0.0037)	(0.0036)	(0.0036)	(0.0054)	(0.0059)	(0.0048)	(0.0049)	(0.0039)
μ_Y	-2.2454	-0.8565	-0.2950	-0.2678	-1.1056	-1.1416	-0.7375	-0.5470	-0.7184	-0.3839	-1.3081
	(1.2624)	(0.7052)	(0.5911)	(0.5820)	(0.5399)	(0.4557)	(0.6223)	(0.6589)	(0.7225)	(0.6482)	(0.7112)
σ_Y	2.0853	1.9256	1.7151	1.7082	1.6005	1.5947	1.6166	1.6394	1.6958	1.6951	1.6871
	(0.4007)	(0.3319)	(0.2539)	(0.2541)	(0.1853)	(0.1746)	(0.1965)	(0.2167)	(0.2432)	(0.2440)	(0.2280)
ρ_J	-0.7342	-2.2089	-2.8920	-2.8966	-0.2434	-0.1429	-1.0584	-1.5945	-1.7742	-2.3316	-1.0089
	(0.7870)	(0.7114)	(0.8380)	(0.7983)	(0.9804)	(0.7495)	(1.4293)	(1.5380)	(1.4725)	(1.2764)	(0.9187)
μ_V	1.7509	1.0941	0.7296	0.7129	0.5586	0.5682	0.5745	0.6102	0.6851	0.6798	0.6085
	(0.5201)	(0.2842)	(0.1705)	(0.1672)	(0.1031)	(0.1020)	(0.1320)	(0.1522)	(0.1764)	(0.1723)	(0.1203)
δ	–	–	–	–	–	–	–	–	–	–	-0.1710
	–	–	–	–	–	–	–	–	–	–	(0.0667)

First, the Cev-diffusion parameter β_3 in SV models is estimated to be slightly higher than one. This confirms results in Christoffersen et al. (2010), Kaeck and Alexander (2013) and others who argue in favor of non-affine diffusion dynamics. Secondly, in linear drift specifications the speed of mean reversion α_2 differs substantially depending on whether the diffusion is affine or not. We obtain the highest value for the Sqr diffusion model with an estimate of -0.021, whereas the value in the 3/2 specification drops to -0.0006. And thirdly, we find a strong leverage effect, with posterior means for ρ close to -0.60 in all models. These results deviate slightly from the estimates in Eraker et al. (2003), who report correlation coefficients between -0.40 and -0.48. Our results in jump-augmented stochastic volatility models in Table 3 confirm that jumps in AffineSqr are exclusively capturing the most extreme events, with a low jump probability $\lambda = 0.0044$ and an average jump size of $\mu_Y = -2.8786$. Jumps in non-affine specifications occur more frequently, for instance the highest jump frequency is obtained for the Poly3/2 model with a daily jump probability of 0.0175. The impact of jumps in the non-affine models is however often less severe, with average jump sizes of approximately -1% as can be seen in Table 3 and 4.

Panel A: SV Affine Drift



Panel B: SV Poly Drift

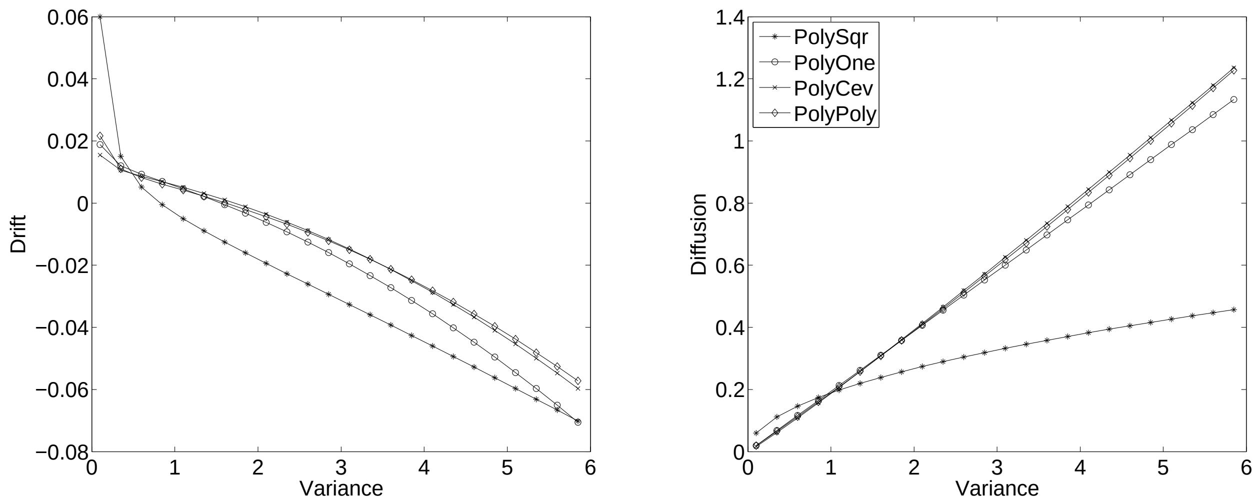


Figure 1: Drift and Diffusion Functions for Various Pure SV Model Specifications

The graphs show the variance drift functions and variance diffusion functions of the AffineSqr, AffineOne, AffineCev and AffinePoly models and the PolySqr, PolyOne, PolyCev and PolyPoly. The variance drift and diffusion functions are nested in the general specifications given by $\alpha_0 + \alpha_1 \frac{1}{V_t} + \alpha_2 V_t + \alpha_3 V_t^2$ and $\beta_0 + \beta_1 V_t + \beta_2 V_t^{\beta_3}$, respectively. For more details on the general stochastic variance specification see Section 2. The plots are constructed as follows: In each step of the MCMC algorithm, we use the structural parameters to calculate the drift and diffusion function for variance levels between 0.1 to 6 (in increments of 0.25). The plots depict the average values for the drift and diffusion function over all iterations. The data set used for estimating the posterior means consists of log returns of the S&P index from January 1983 to December 2013. The returns are measured as daily percentage log returns, so that a daily variance of 1 yields a yearly standard deviation of $\sqrt{252} = 15.87\%$.

To show the difference between the estimated specifications, in Figure 1 we plot the functions α and β evaluated at the posterior means for different models in the SV class.[15] The top two graphs highlight the different drift and diffusion behavior for models with an affine drift. It is apparent that the Sqr model estimates a significantly lower diffusion level during high volatility regimes. Models with more flexibility in the diffusion term, such as Poly and Cev, produce functions that remain relatively close to the One specification. Interestingly, the choice of the diffusive term also has a marked effect on the drift function. The square-root diffusion models in particular require a substantially stronger pull towards their long-term variance level, whereas the mean reversion in more flexible specifications is less pronounced. This can be seen by noting that the drift function of the square root diffusion models exhibits a steeper slope compared to the other model specifications. A possible explanation for this finding is that with a stronger mean reversion, the affine model can generate large negative variance increments during periods of high volatility. The remaining two graphs in Figure 1 show drift and diffusion functions for models with non-affine drift specifications. There are two noteworthy findings. First, the diffusion function remains almost unaltered when extending the affine to a more flexible drift specification. And secondly, the Sqr specification again provides the most extreme behavior, especially around the zero boundary. Finally, the Sqr model provides the lowest estimate for the average variance level that is given by the root of the drift function.

Turning to the Box-Cox models we note that the parameters α_0, α_2, and β_0 are not directly comparable to the remaining variance specifications since they relate to the Box-Cox transformed process H_t. Concerning the jump parameters, we find similar results to those discussed earlier. Jumps are rare at about 2.5 jumps per year and on average negative. Also, the leverage effect is pronounced with estimates for ρ between -0.59 and -0.70. Most

[15]The functions for SVJ and SVCJ follows similar patterns and are not reported for brevity. Details are available on request.

importantly, the parameter δ is estimated at around -0.2 for all model classes. As shown in Yu et al. (2006) the Box-Cox transformed model specification reduces to well-known stochastic volatility models for certain parameter combinations. In particular, for $\delta = 0$ the model reduces to the log volatility model analyzed in Yu (2005). Our estimation results clearly reject this special case which confirms the results in Zhang and King (2008).

4.2 Return Fit

To compare models based on their ability to fit return observations over the whole sample, we report the DIC statistic in Table 5, as well as the model complexity penalty term p_D and the model fit term $\bar{D}$ for all competing specifications. Overall we find that p_D and $\bar{D}$ exhibit expected patterns. For the model complexity p_D, we report high statistics for the most complex SVCJ models and values decline for SVJ and SV specifications. The model fit component $\bar{D}$ is lowest, indicating best model fit, for the most complex models. Table 5 provides several interesting results.

First, we confirm findings in Ignatieva et al. (2015) that jumps improve the in-sample performance, as models in the SV class consistently exhibit the highest DIC values,[16] whereas SVCJ specifications, except for AffineSqr, outperform SVJ. Secondly, non-affine diffusion models in general perform considerably better than affine models. For all model classes, the AffineSqr specification is outperformed by several non-affine specifications. Thirdly, no clear preferred drift and diffusion specification arises across the three model classes. Within the SVCJ class the Affine3/2 specification ranks top, whereas in the SVJ and the SV classes, respectively, PolyOne and PolySqr perform best. We also observe an interesting structural relationship between the jump model complexity and the impact of flexible drift and diffusion

[16] The lower the DIC statistic the better is the performance of the model.

Table 5: Rankings of Models by DIC

The table shows the Deviance Information Criterion (DIC) rankings for all models based on S&P 500 data for the time period January 1983 until December 2013. The DIC column gives the overall DIC value where lower values indicate a better model performance. The Deviance Information Criterion consists of two parts p_D the penalty term measuring model complexity and $\bar{\text{D}}$ measuring model fit (see Spiegelhalter et al., 2002).

Model	**DIC**	$\mathbf{p}_D$	$\bar{\mathbf{D}}$
SVCJ Affine3/2	15133.5	4448.8	10684.8
SVCJ Poly3/2	15244.2	4626.8	10617.4
SVCJ AffineCev	15835.8	4415.8	11420.0
SVCJ PolyCev	16230.9	4255.1	11975.8
SVCJ AffinePoly	16914.2	4160.8	12753.4
SVCJ PolyPoly	17368.5	3960.6	13407.9
SVCJ PolyOne	17649.5	3827.3	13822.2
SVCJ AffineOne	17711.3	3794.9	13916.4
SVCJ PolySqr	18034.0	3561.2	14472.8
SVCJ BoxCox	18222.9	3958.9	14264.0
SVJ PolyOne	18240.1	3454.2	14786.0
SVJ AffineOne	18246.7	3446.3	14800.4
SVJ PolySqr	18254.5	3379.6	14874.9
SVJ AffinePoly	18278.0	3631.4	14646.6
SVCJ AffineSqr	18285.4	3436.3	14849.1
SVJ AffineCev	18285.6	3534.1	14751.5
SVJ PolyCev	18302.2	3488.1	14814.1
SVJ PolyPoly	18321.5	3507.0	14814.4
SVJ AffineSqr	18485.2	3245.4	15239.8
SVJ BoxCox	18577.3	3773.9	14803.4
SVJ Affine3/2	18691.6	3437.3	15254.3
SVJ Poly3/2	18735.2	3390.1	15345.1
SV PolySqr	18956.3	2771.9	16184.5
SV PolyOne	18980.4	2802.8	16177.6
SV AffineOne	18987.3	2790.4	16196.9
SV AffinePoly	18990.3	2833.7	16156.6
SV PolyCev	19024.8	2813.8	16211.0
SV AffineCev	19038.8	2819.6	16219.2
SV PolyPoly	19064.2	2873.7	16190.6
SV AffineSqr	19119.0	2656.5	16462.6
SV Affine3/2	19432.8	2691.1	16741.7
SV Poly3/2	19455.7	2660.3	16795.4
SV BoxCox	19521.4	2834.6	16686.8

specifications. Within the SVCJ class, the model ranking is driven exclusively by the diffusion component. This can be seen from the fact that combinations of a particular diffusion component with the two drift specifications, Affine or Poly, show little difference in their fit statistics. We find that the ranking is stable with respect to the diffusion models pairs, Affine3/2 and Poly3/2 for instance outperform both AffineCev and PolyCev. This structure can still be observed to some extent for SVJ models, but disappears for the SV model class. Therefore, for the most complex SVCJ model class the diffusion component seems to be able clearly identify model performance. We return to these results when discussing out-of-sample model performance below. Finally, we find that the Box-Cox variance models rank last in the SVCJ and SV model class and third to last in the SVJ model class. This shows that modeling non-linearities directly in the variance process results in a superior model performance compared to the non-linear Box-Cox transformation.

To focus on our main findings, we confine the discussion in the following to the three best performing models as well as the affine benchmark model in SV, SVJ and SVCJ.[17] First, to test the stability of the model ranking over time, we re-estimate the models during subsamples and apply two distinct setups. First, we start with a sample from 1983 to 1999 and then subsequently add one more year to update parameter estimations and model rankings. In the second setup we divide the whole sample period into three subsets, 1983 to 1993, 1994 to 2003 and 2004 to 2013, and study parameter estimates and model rankings for these non-overlapping subsets. The two exercises provide distinct robustness checks. The first setup addresses the question of how the model ranking changes if the information set increases over time, whereas the second setup looks at three distinct non-overlapping time periods and compares the model ranking between those mutual exclusive subsets. While computationally intensive, this procedure provides further insights as far as the robustness of the findings in Section 4.2 are concerned.

[17] Result for the full set of models are available from the author upon request.

In Table 6, we present model rankings for the fifteen increasing sub-samples. Overall, these results confirm our previous findings. In particular, jump specifications outperform pure stochastic variance models (as indicated by the *Min* column in Table 6); the best minimum rank of a SV model is 9, implying that all 8 other jump-diffusion models rank higher in all sub-periods. In addition, non-affine diffusion models outperform their affine counterpart in all sub-samples. Comparing the ranking of the AffineSqr specification with the ranking of the non-affine specifications within each model class, we find that the AffineSqr specification is dominated by the non-affine specification in each sub-sample. In the SVCJ class, the Affine3/2 specification has the best average ranking (1.80, see the *Mean* column in Table 6) and for the SVJ and the SV class, the AffineOne (4.33) and PolyOne (9.4) show the best average performance, respectively. This provides further evidence that no single combination of drift and diffusion function results in superior performance across SV, SVJ and SVCJ.

Table 6: Rankings of Models by DIC (Sub-samples)

The table shows the Deviance Information Criterion (DIC) rankings for all models based on S&P 500 data for different time periods. Each column ranks all models for a data set starting in January 1983 and ending in December of the year indicated in the column header. The best DIC ranking is indicated by a (1) and the worst is indicated by (12). The last two columns give the minimum ranking and average ranking of a model over all data periods.

model	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**Min**	**Mean**
SV AffineSqr	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12.00
SV PolySqr	9	11	11	11	11	9	11	11	10	10	10	9	9	9	9	9	10.00
SV AffineOne	11	10	10	10	10	11	10	10	11	11	11	11	11	11	11	10	10.60
SV PolyOne	10	9	9	9	9	10	9	9	9	9	9	10	10	10	10	9	9.40
SVJ AffineSqr	7	8	8	8	8	8	8	8	8	8	8	8	8	8	8	7	7.93
SVJ PolySqr	6	6	6	7	7	7	6	7	6	7	7	6	6	5	6	5	6.33
SVJ AffineOne	5	4	4	4	5	4	4	4	4	5	4	4	4	6	4	4	4.33
SVJ PolyOne	4	5	5	5	4	5	5	5	5	4	5	5	5	4	5	4	4.73
SVCJ AffineSqr	8	7	7	6	6	6	7	6	7	6	6	7	7	7	7	6	6.67
SVCJ Affine3/2	2	1	1	3	1	2	3	3	2	3	1	1	1	2	1	1	1.80
SVCJ Poly3/2	1	2	2	1	3	3	2	1	3	1	2	3	3	3	3	1	2.20
SVCJ AffineCev	3	3	3	2	2	1	1	2	1	2	3	2	2	1	2	1	2.00

Figure 2 shows the evolution of posterior means for all structural parameters of the SVCJ PolyCev model over the different sub-samples. Overall, we find that parameters remain relatively stable over time and posterior means for the full sample are often within the 90% confidence sets estimated from the shortest sample. The variation of parameter estimates is slightly larger for the drift parameters α_0 to α_3, whereas jump parameters remain stable over time. These results also highlight a downward trend for the correlation between returns and the variance process as ρ-estimates decrease around the beginning of the century (dot-com bubble and its subsequent burst), results that explain the slightly more negative estimate in Tables 2 to 4 compared to empirical results in Eraker et al. (2003).

In Table 7 we report fit results for mutually exclusive datasets as well as summary statistics for the different data periods. The three periods show slightly different volatility estimates, the most noticeable difference, however, is the very large kurtosis for the period from 1983 to 1993. This large kurtosis is fully driven by one extreme return of -23% on Oct 19, 1987. Without this outlier, the different data periods show overall similar return characteristics. With respect to the fit statistics, we find similar patterns to those discussed above. First, during all three samples the pure SV models are outperformed by SVJ and SVCJ models. For the first two samples, the SV models are strictly outperformed, whereas in the final sample two non-affine SV models perform better than the affine SVJ specification. Second, we find that in each jump class, the AffineSqr specification is inferior to non-affine specifications. Within the SVCJ class the best performing model over all three periods is the SVCJ Affine3/2 model as indicated by the lowest mean value of 1.33. This is similar to the results for the previous sub-set analysis where SVCJ Affine3/2 was also performing best. Overall, this indicates that model choice is affected by the complexity of the jump distribution.

Table 7: Rankings of Models by DIC (Sub-samples)

The table shows the Deviance Information Criterion (DIC) rankings for all models based on S&P 500 data for different time periods. Column two to four rank all models for data sets based on time periods 1983 to 1993, 1994 to 2003, and 2004 to 2013, respectively, as indicated in the column headers. The best DIC ranking is indicated by a (1) and the worst is indicated by (12). The last two columns give the minimum ranking and average ranking of a model over all data periods.

Model	**1983 - 1993**	**1994 - 2003**	**2004 - 2013**	**Min**	**Mean**
SV AffineSqr	12	12	12	12	12.00
SV PolySqr	9	11	8	8	9.33
SV AffineOne	11	9	9	9	9.67
SV PolyOne	10	10	11	10	10.33
SVJ AffineSqr	8	8	10	8	8.67
SVJ PolySqr	6	6	5	5	5.67
SVJ AffineOne	7	4	6	4	5.67
SVJ PolyOne	5	5	7	5	5.67
SVCJ AffineSqr	4	7	4	4	5.00
SVCJ Affine3/2	1	1	2	1	1.33
SVCJ Poly3/2	2	2	3	2	2.33
SVCJ PolyCev	3	3	1	1	2.33
No. Obs	2782	2519	2517	–	–
Annual Volatility	0.16	0.18	0.21	–	–
Min	-0.23	-0.07	-0.09	–	–
Max	0.09	0.06	0.10	–	–
1% Percentile	-0.02	-0.03	-0.04	–	–
99% Percentile	0.02	0.03	0.04	–	–
Skewness	-4.34	-0.11	-0.33	–	–
Kurtosis	100.4	6.19	14.03	–	–

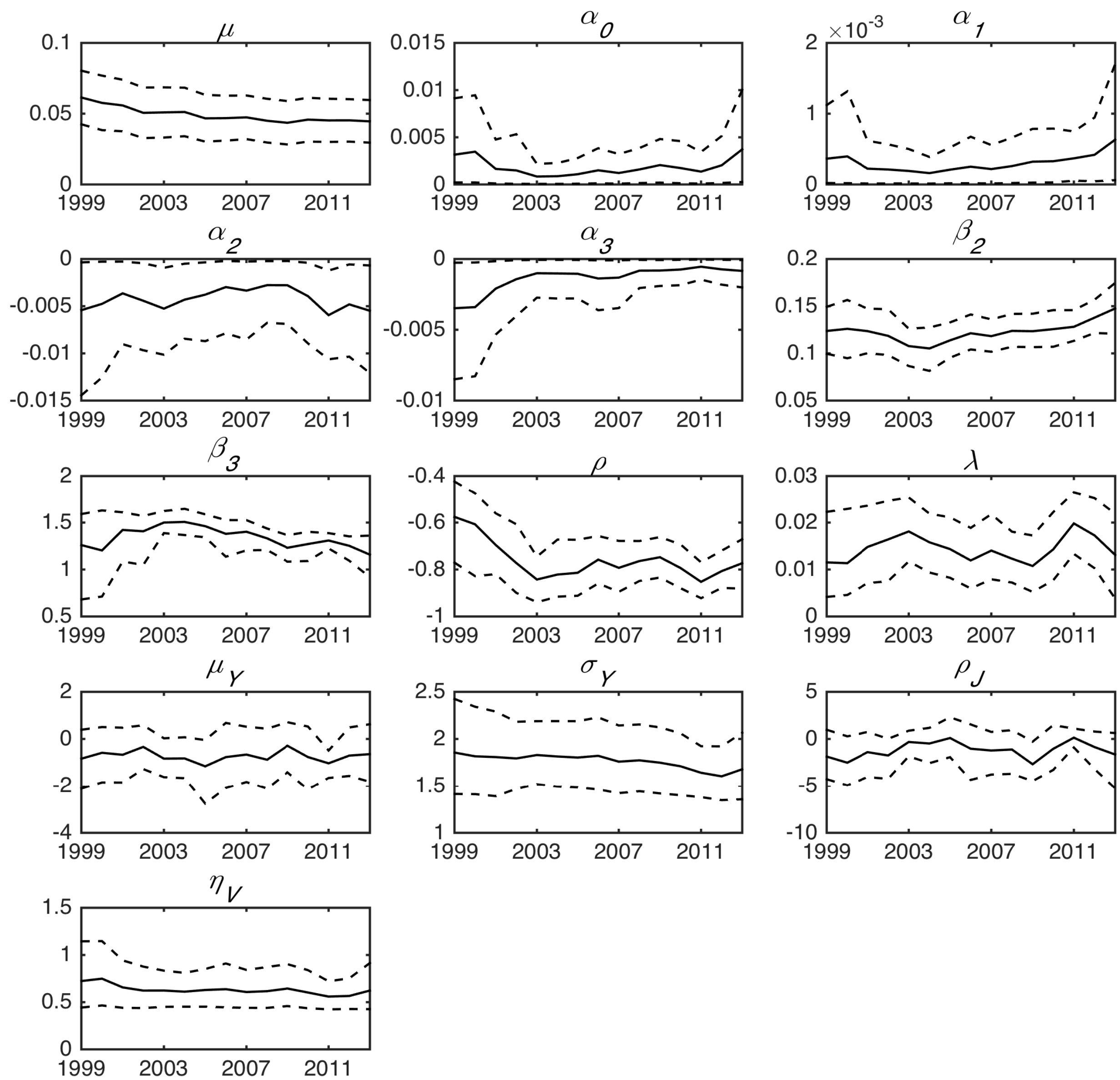


Figure 2: SVCJ PolyCev Model Robustness of Parameter Estimates over Time

The graph shows posterior means for all structural parameters for increasing datasets (solid lines). The first parameter estimates are based on S&P 500 index returns from 1983 to 1999 and we subsequently add one year of data such that the last set of parameters is based on data from 1983 to 2013. The dashed lines depict 90% confidence sets.

4.3 Variance Paths

Visual inspection of estimated variance paths provides valuable information on whether alternative models portray realistic dynamics for the latent state processes. Model comparison could rely on either smoothed variance estimates which are a by-product of the MCMC algorithm but use the entire return data set for variance estimation, or on filtered variance which ensures that variances are estimated using past returns only and by that use the same information set as model free realized variance estimators. We therefore rely on the filtered variance estimator for our analysis.[18] In a first step we focus on how filtered variance estimates differ from popular measures of realized variance calculated from intra-daily returns. We are particularly interested in whether different combinations of drift and diffusion specifications produce distinct patterns in variance time series. Driven by the availability of realized variance estimates, provided by the Oxford Man Institute, we restrict the sample to 2000 to 2013 for this exercise.[19] In particular, we select the realized variance estimator (5-minute using 1-minute subsamples) from the Oxford Man institute.[20] In addition, we obtain the price series for the SPY exchange traded fund from the Center of Research in Security Prices (CRSP) to calculate overnight returns. We then follow the approach proposed in Andersen et al. (2015) and add the squared overnight returns to the realized intra-daily variance to approximate close-to-close integrated variance.

To generate filtered variance paths, we use model parameters estimated on daily S&P 500 returns from 1983 to 1999 and then filter latent state variables during an out-of-sample period from 2000 to 2013. This procedure guarantees that variances are not estimated using future return data. We apply an extension of the continuous-time particle filter described in

[18] In an earlier version of this paper we used a smoothed variance estimator. Results are structurally very similar.

[19] Realized variance estimators are downloaded from the homepage of Oxford Man Institute: http://realized.oxford-man.ox.ac.uk/.

[20] For details on the computation of the realized variance see http://http://realized.oxford-man.ox.ac.uk/documentation/econometric-methods/.

Johannes et al. (2009). Due to space restrictions, we limit our discussion in this section to the SVCJ model class, since results for other model classes are structurally similar. Furthermore, we do not detail the Cev specification, as this model does not provide notable differences to the 3/2 models.

Figures 3 and 4 provide filtered variance paths and realized variances during the high-volatility period of the Lehman default in 2008 and during a low-volatility regime at the end of 2005 (labeled *calm period*). To gage the effect of the drift specification, we compare affine and polynomial drift specifications in each row, whereas columns are used for the comparison of alternative diffusion specifications. Figure 3 shows that the drift specification has a very limited effect on estimated variance paths. For example, the first row compares AffineSqr with PolySqr, and differences in the estimated paths are only very minor. This finding is very robust and holds for all Affine and Poly drift specification. Furthermore, this graph confirms the superiority of non-affine variance specifications in generating rapid movements during high volatility regimes. This is particularly evident in a row-wise comparison of different diffusion specifications. For higher values of the parameter β_3, estimated variance paths provide a much better fit to the realized path. This holds particularly true for models with a β_3 that is greater then one. On the other hand, Figure 4 shows that the variance paths during the low volatility regime appear indistinguishable for different models and can therefore be captured by all specifications. This shows that the differences in model specifications only become visible during times of market stress. In other words, it is modeling stressed market scenarios in particular that calls for more sophisticated diffusion specifications.

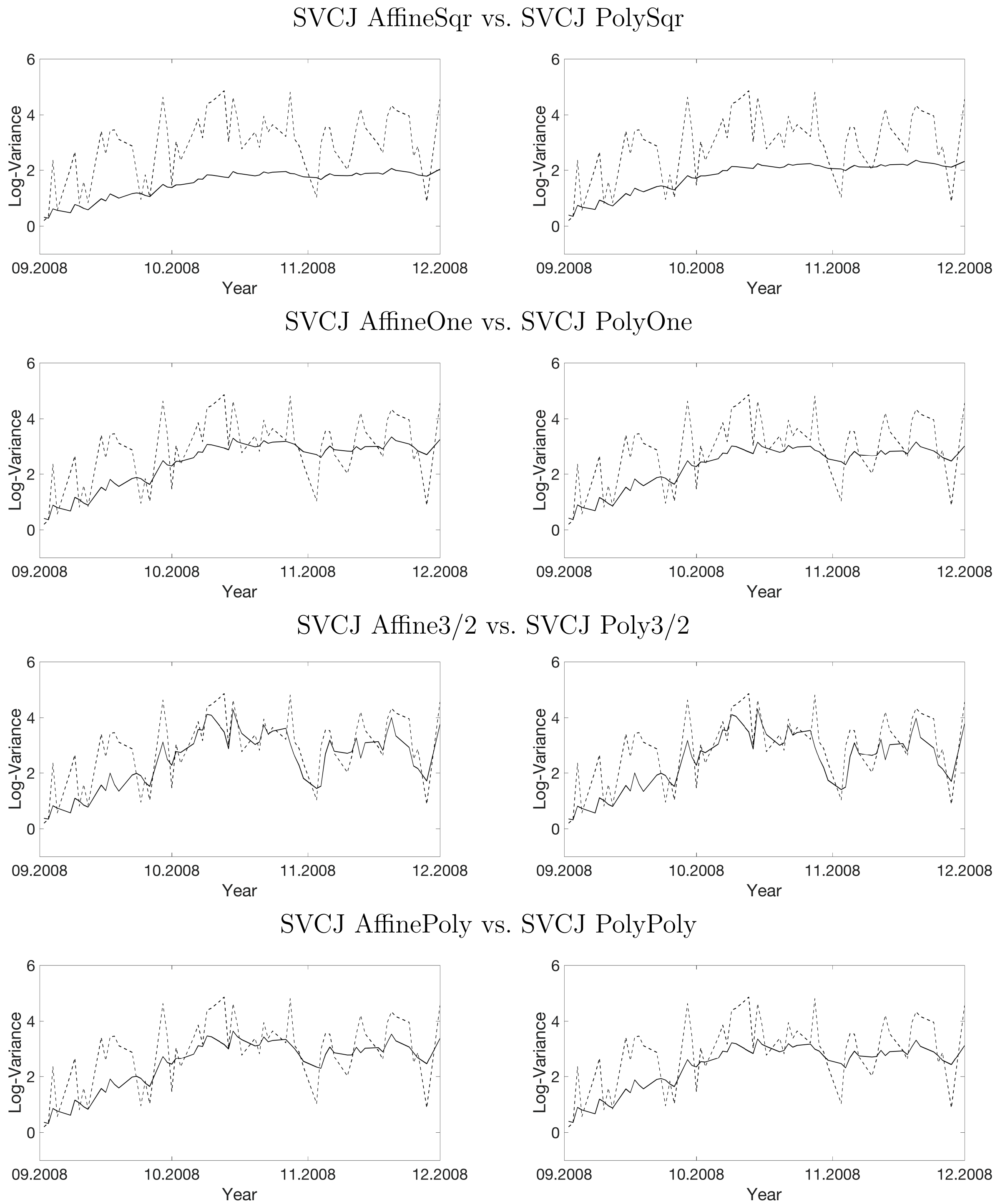


Figure 3: RV vs. Filtered Variances - Lehman Crisis

This figure shows filtered variances (solid lines) for the AffineSqr, AffineOne, Affine3/2, AffinePoly, and PolySqr, PolyOne, Poly3/2 and PolyPoly SVCJ models generated by a particle filter algorithm (see Johannes et al. (2009)) against realized variances (dashed lines). Data period is the high volatility regime around the Lehman crisis period from September 2 to December 31 2009. Realized variance estimator (5-minute using 1-minute subsamples) are downloaded from the website of the Oxford Man Institute (http://realized.oxford-man.ox.ac.uk/).

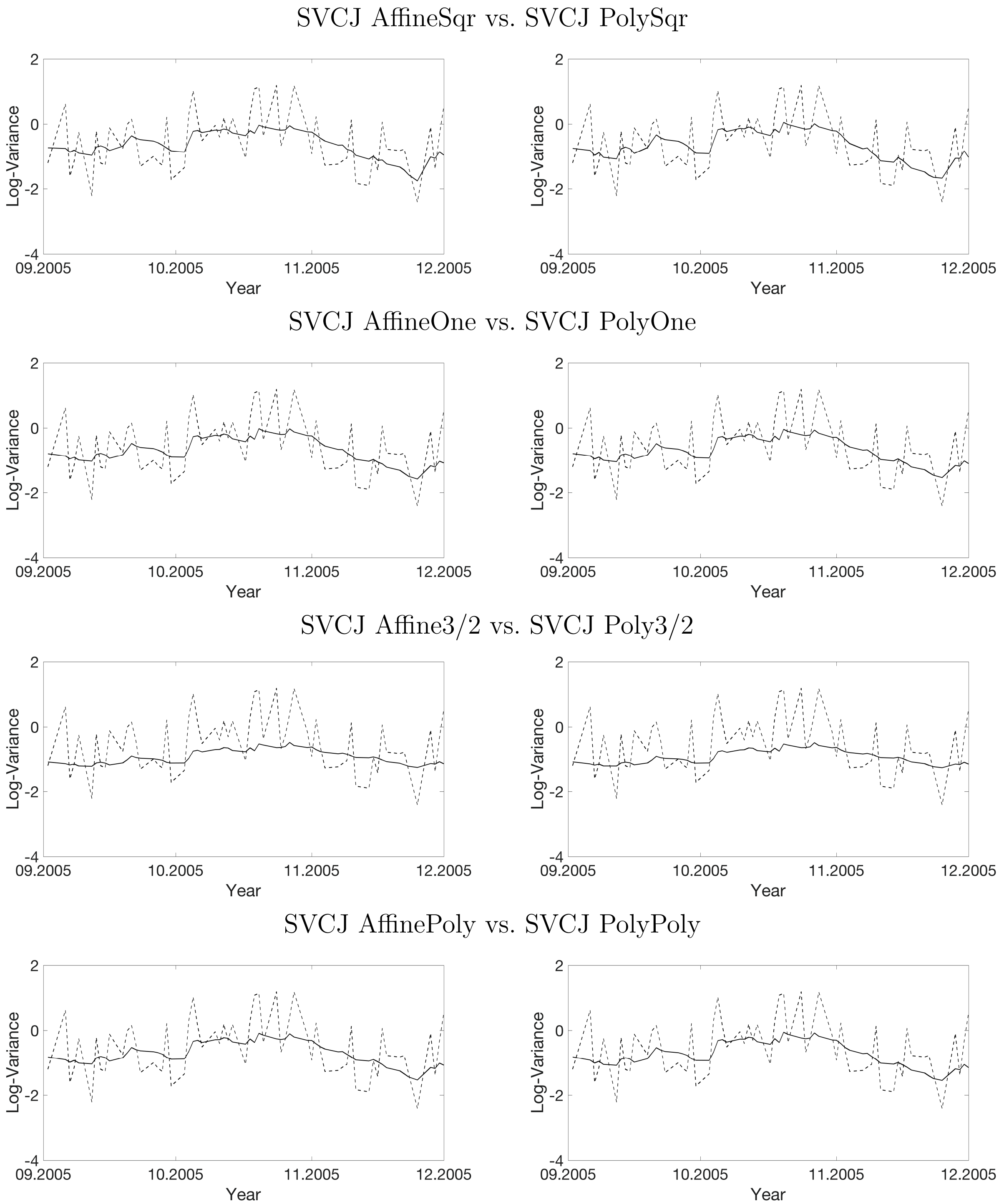


Figure 4: RV vs. Filtered Variances - Calm Period

This figure shows filtered variances (solid lines) for the AffineSqr, AffineOne, Affine3/2, AffinePoly, and PolySqr, PolyOne, Poly3/2 and PolyPoly SVCJ models generated by a particle filter algorithm (see Johannes et al. (2009)) against realized variances (dashed lines). Data period is a calm market period September 1 to December 30 2005. Realized variance estimator (5-minute using 1-minute subsamples) are downloaded from the website of the Oxford Man Institute (`http://realized.oxford-man.ox.ac.uk/`).

4.4 Realized Variance Regressions

For a more rigorous comparison of realized and model-based variances, we follow Hansen and Lunde (2006) and study the following regression equation for all competing models $m \in \mathcal{M}$:

$$\log(RV_t) = a_m + b_m \log(\hat{V}_t^{F,m}) + \varepsilon_{t,m} \tag{13}$$

where RV_t and $\hat{V}_t^{F,m}$ respectively denote the realized and filtered variance on day t.[21] We include the return between $t-1$ and t in the filtering of the variance path and compare the day-t estimate to the realized variance for day t. Our realized variance measure (constructed as described in the last section) includes squared overnight returns to capture the close-to-close integrated variance. Note, that our realized variance estimator is driven by a continuous variation component and a jump component. Alternatively, we could have used the bipower variation estimator (see Barndorff-Nielsen and Shephard, 2004) which only captures the diffusive component of the log-price variation. As jumps are rare and since our filtered variance is based on close-to-close returns, we userealized variance but have confirmed that our conclusions are unchanged if we use bipower variation (untabulated). Note that our model-based results are purely out-of-sample, as our filtering algorithm is based on parameter estimates from the 1983-1999 return data set. We estimate regression Equation (13) by Bayesian statistical methods with diffuse priors and provide results in Tables 8 and 9.

A well-specified variance model should generate estimates of a_m close to zero, b_m close to one, and a relatively high R^2.[22] Table 8 shows the results for a sub-sample from September 1 to December 30, 2005. This time period is characterized by low market volatility. We observe

[21]We also ran the analysis in levels of RV_t and $\hat{V}_t^{F,m}$ which yields structurally the same results. The log-specifications is the preferred one in Hansen and Lunde (2006), though, and is consistent with our analysis in Section 5.

[22]As RV includes jumps, the empirical estimates of a_m and b_m may deviate from zero and one, even in a correctly specified model.

Table 8: Regression of RV on Filtered Variances - Calm Period

This tables provides estimation results for the regression model $RV_t = a_m + b_m \hat{V}_t^{F,m} + \varepsilon_{t,m}$, where RV_t denotes the daily realized variances and $\hat{V}_t^{F,m}$ denotes the filtered variances of model m on day t. The regression is based on realized and filtered variances for a calm market time period September 1 to December 30 2005. Realized variance estimator (5-minute using 1-minute subsamples) are downloaded from Oxford Man Institute (http://realized.oxford-man.ox.ac.uk/)

	a_m		b_m		R^2	
Models	Mean	S. E.	Mean	S. E.	Mean	S. E.
SV Model Class						
Affine SQR	-0.0044	0.0122	1.1747	0.0386	0.5627	0.0018
Poly SQR	-0.0020	0.0127	1.1525	0.0380	0.5562	0.0011
Affine One	-0.0020	0.0126	1.1445	0.0373	0.5521	0.0013
Poly One	-0.0017	0.0127	1.1192	0.0374	0.5494	0.0013
Affine 3/2	-0.0046	0.0132	1.2898	0.0455	0.5012	0.0018
Poly 3/2	-0.0049	0.0134	1.3174	0.0478	0.5028	0.0018
Affine Cev	-0.0017	0.0128	1.1318	0.0374	0.5424	0.0013
Poly Cev	-0.0021	0.0131	1.1216	0.0392	0.5435	0.0013
Affine Poly	-0.0029	0.0129	1.1568	0.0395	0.5472	0.0012
Poly Poly	-0.0034	0.0132	1.1452	0.0375	0.5410	0.0014
SVJ Model Class						
Affine SQR	-0.0061	0.0122	1.1115	0.0363	0.5417	0.0021
Poly SQR	-0.0032	0.0127	1.1109	0.0366	0.5380	0.0017
Affine One	-0.0045	0.0121	1.1155	0.0363	0.5330	0.0016
Poly One	-0.0043	0.0126	1.1008	0.0379	0.5298	0.0018
Affine 3/2	-0.0072	0.0126	1.3490	0.0480	0.4928	0.0024
Poly 3/2	-0.0052	0.0134	1.3810	0.0488	0.4977	0.0024
Affine Cev	-0.0052	0.0122	1.1488	0.0387	0.5210	0.0018
Poly Cev	-0.0047	0.0129	1.1285	0.0387	0.5283	0.0018
Affine Poly	-0.0049	0.0123	1.1669	0.0403	0.5245	0.0021
Poly Poly	-0.0039	0.0131	1.1540	0.0401	0.5310	0.0019
SVCJ Model Class						
Affine SQR	-0.0049	0.0126	1.1011	0.0375	0.5278	0.0019
Poly SQR	-0.0029	0.0125	1.0543	0.0357	0.5289	0.0015
Affine One	-0.0042	0.0129	1.0582	0.0375	0.5221	0.0019
Poly One	-0.0038	0.0131	1.0613	0.0379	0.5215	0.0016
Affine 3/2	-0.0071	0.0128	1.2896	0.0476	0.4834	0.0031
Poly 3/2	-0.0071	0.0128	1.2823	0.0482	0.4870	0.0029
Affine Cev	-0.0061	0.0129	1.0618	0.0381	0.5018	0.0023
Poly Cev	-0.0053	0.0123	1.0568	0.0381	0.5103	0.0020
Affine Poly	-0.0051	0.0128	1.0564	0.0368	0.5159	0.0020
Poly Poly	-0.0040	0.0131	1.0547	0.0382	0.5162	0.0018

that all models generate estimated values for a_m that are close to zero. As expected, estimated values for b_m are slightly higher than one and show some variation across the different models. The SVCJ PolySqr model is closest to one with an estimated parameter value of 1.0543 (which is not significantly different from one), whereas the SVJ Poly3/2 model differs the most from one with $b_m = 1.3810$ (which is significantly different from one). In terms of the R^2-values we find variation across models with values between 0.48 and 0.57. That is, based on regression results for the low-volatility period we can confirm the result in Section 4.3 and conclude that models show minor differences for the calm period although the non-affine jumps models are slightly more successful in capturing realized variance dynamics. It is interesting to note that the AffineSqr specifications perform very well in terms of R^2 for periods characterized by low market volatility.

Table 9 shows estimation results for the high volatility period around the Lehman crisis. In contrast to earlier results, we observe large differences between affine and non-affine model specifications for both b_m and R^2. All models with a Sqr diffusion exhibit large b_m values (between 1.3 and 1.7). In addition, these models also exhibit the lowest R^2 values (between 0.82 and 0.84). These statistics are significantly improved by models with a diffusion parameter β_3 greater than 1. We find that, across all model classes, the R^2 increases to over 0.9 for specifications with β_3 greater than 1. Further, we observe that the coefficient b_m is much closer to one, reaching values of 1.07. Finally, altering the drift specifications for identical diffusion models leads to minor improvements compared to changes to the diffusion setup. For example, for the SVJ model class moving from the AffineSqr to the PolySqr specification changes the R^2 from 0.8272 to 0.8278, and b_m from 1.4889 to 1.3629. If we fix the drift to the Affine specification and alter the diffusion to the 3/2 model setup we observe an increase of the R^2 to 0.9162 and the slope parameter b_m decreases to 1.0913. This suggests that model improvements are mainly driven by the diffusion specification.[23]

[23] Hansen and Lunde (2006) propose to either use a log-on-log or a level-on-level model for regression equation

In summary, realized variance regressions confirm our earlier findings in Section 4.3. First, non-affine specifications are able to mimic realized variances more closely than their affine counterparts. Secondly, the differences stem mainly from the diffusion setup. And thirdly, the differences are most extreme during crisis periods and vanish almost completely during calm market regimes.

4.5 Implications for Variance Swaps

In this section, we address implications of our results for assessing variance risk premia. Variance contracts have gained considerable attention in the finance industry and academia over the last decades. Active areas of research include: developing replication strategies, e.g., Britten-jones and Neuberger (2000), the pricing of variance risk, e.g., Coval and Shumway (2001), or specifying general equilibrium models that are able to explain the variance risk premium, e.g., Drechsler and Yaron (2011).

In our analysis, we have an investor in mind who wants to use a variance model to determine the expected return of an investment into variance swaps. Variance swaps allow investors to hedge the variance risk of some underlying asset such as interest rates or stock indices. By entering the swap contract the investor pays a fixed amount and receives a floating value based on the realized variance of the underlying, or vice versa. The payoff of a variance swap is the net difference between the two payments. This difference is also called the variance risk premium and has attracted considerable attention in the literature (see Carr and Wu, 2009). Our model classification allows us to assess the ability of models to generate realistic expectations of variance risk premia and to gage the model risk related to using alternative

(13). The log-on-log specification is less sensitive to outliers than a level-on-level setup. In unreported results we estimate the same relation using levels. The conclusions are qualitatively equal to our presentation here. However, quantitatively the differences across models in terms of R^2 and values for b_m for the crisis period are much larger. This is to be expected, since the financial crisis generated large variance spikes that can be characterized as outliers.

Table 9: Regression of RV on Filtered Variances - Crisis Period

This tables provides estimation results for the regression model $RV_t = a_m + b_m \hat{V}_t^{F,m} + \varepsilon_{t,m}$, where RV_t denotes the daily realized variances and $\hat{V}_t^{F,m}$ denotes the filtered variances of model m on day t. The regression is based on realized and filtered variances for the Lehman crisis time period September 2 to December 31 2008. Realized variance estimator (5-minute using 1-minute subsamples) are downloaded from Oxford Man Institute (http://realized.oxford-man.ox.ac.uk/)

	a_m		b_m		R^2	
Models	Mean	S. E.	Mean	S. E.	Mean	S. E.
			SV Model Class			
Affine Sqr	0.0032	0.0104	1.4454	0.0156	0.8424	0.0003
Poly Sqr	0.0048	0.0107	1.3139	0.0147	0.8370	0.0003
Affine One	0.0024	0.0103	1.0838	0.0098	0.8929	0.0002
Poly One	0.0023	0.0100	1.1753	0.0110	0.8971	0.0001
Affine 3/2	0.0040	0.0098	1.0747	0.0091	0.9246	0.0002
Poly 3/2	0.0038	0.0102	1.0926	0.0096	0.9220	0.0003
Affine Cev	0.0022	0.0101	1.0663	0.0099	0.9012	0.0002
Poly Cev	0.0018	0.0099	1.1594	0.0104	0.9042	0.0001
Affine Poly	0.0024	0.0098	1.0847	0.0095	0.9082	0.0002
Poly Poly	0.0018	0.0100	1.1525	0.0105	0.9104	0.0001
			SVJ Model Class			
Affine Sqr	0.0048	0.0110	1.4889	0.0165	0.8272	0.0004
Poly Sqr	0.0051	0.0113	1.3629	0.0157	0.8278	0.0004
Affine One	0.0034	0.0106	1.0851	0.0107	0.8795	0.0002
Poly One	0.0022	0.0101	1.1675	0.0113	0.8884	0.0002
Affine 3/2	0.0044	0.0097	1.0913	0.0094	0.9162	0.0003
Poly 3/2	0.0044	0.0100	1.1011	0.0096	0.9159	0.0003
Affine Cev	0.0026	0.0097	1.0678	0.0100	0.8972	0.0002
Poly Cev	0.0019	0.0100	1.1412	0.0106	0.8985	0.0002
Affine Poly	0.0027	0.0097	1.0831	0.0098	0.9059	0.0002
Poly Poly	0.0022	0.0098	1.1508	0.0103	0.9079	0.0002
			SVCJ Model Class			
Affine Sqr	0.0051	0.0104	1.7085	0.0185	0.8420	0.0003
Poly Sqr	0.0056	0.0111	1.4604	0.0168	0.8333	0.0004
Affine One	0.0044	0.0102	1.1198	0.0116	0.8651	0.0003
Poly One	0.0030	0.0105	1.1833	0.0118	0.8736	0.0003
Affine 3/2	0.0088	0.0106	1.1506	0.0119	0.8781	0.0006
Poly 3/2	0.0074	0.0097	1.1557	0.0107	0.8853	0.0005
Affine Cev	0.0045	0.0098	1.1015	0.0101	0.8880	0.0003
Poly Cev	0.0041	0.0104	1.1596	0.0117	0.8851	0.0003
Affine Poly	0.0041	0.0099	1.1135	0.0104	0.8840	0.0003
Poly Poly	0.0039	0.0104	1.1582	0.0113	0.8787	0.0003

specifications.

The literature has proposed various alternative definitions of realized variance in the context of variance risk premium calculations. It is common, for instance, to use the sum of squared returns or squared log returns over a trading month. Neuberger (2012) and Bondarenko (2014), however, show that an alternative definition based on the difference between simple and log returns has theoretical advantages for variance risk premium calculations. In this section, we use the measure proposed in Neuberger (2012) and Bondarenko (2014) which is given by

$$\widetilde{RV}_t = 2\sum_{i=1}^{N_t}(r_{t,i} - \log(1 + r_{t,i})) \tag{14}$$

where $r_{t,i}$ denotes the i-th daily simple return in month t and N_t is the number of trading days in month t. This measure has three advantages compared to using the sum of squared (log) returns. First, the risk-neutral expectation of this measure can be inferred from option prices for (risk-neutral) martingale processes and this expectation is identical to the (properly scaled) squared VIX index. Second, the risk-neutral expectation is model-free, in particular it accounts for jumps in the asset price process. Finally, the risk-neutral expectation is independent of the sampling frequency of the returns used in the realized variance definition.[24] All of our empirical results below are robust to using alternative definitions.

We follow Carr and Wu (2007) and Bondarenko (2014) and define the expected variance risk premium as follows

$$\log\left(E_{t-1}\left(\widetilde{RV}_t\right)\right) - \log(VIX_{t-1}^2) \tag{15}$$

[24] For our application, the only small bias arises due to the fact that we simulate S&P 500 index returns and not futures returns. Because we base our simulations on monthly realized variances, this bias is likely to be negligible. We have also used other definitions of realized variance and find qualitatively and quantitatively similar results.

where the first term denotes the logarithm of the expected realized variance under the real-world probability measure and the second term is the logarithm of the squared VIX index, scaled to a monthly frequency. This definition effectively compares the risk-neutral and the real-world expectation of monthly realized variance, that is $VIX_{t-1}^2 = E_{t-1}^*\left(\widetilde{RV}_t\right)$ where E^* denotes the expectation under the risk-neutral measure.

The computation of the expected realized variance in Equation (15) depends on a parametric variance model, and we compare alternative specifications analyzed in this paper to calculate this expectation. The expectation in Equation (15) is approximated by simulation using our estimation and filtering output, based on 22 trading days. In addition we compute the ex-post realization of this investment as

$$\log\left(2\sum_{i=1}^{N_t}(r_{t,i} - \log(1 + r_{t,i}))\right) - \log(VIX_{t-1}^2), \tag{16}$$

where we use the simple return of the S&P 500 index over one trading day for $r_{t,i}$. Cose-to-close returns for the S&P 500 index are downloaded from CRSP and the VIX index is from the CBOE website. Note that we calculate monthly expected variance risk premia from 2000-2013 to avoid overlapping periods.

The results for this exercise are summarized in Table 10 and Figure 5. First, consistent with previous literature, integrated variance under the risk-neutral measure is higher than under the real world measure, resulting in negative variance risk premia. The mean ex-post return realization is -0.46. In general we observe that models from the SV and SVJ class overestimate (less negative) the ex-post realization of the variance premium. The only exception is the SV and SVJ 3/2 specification, which provides average values similar to the ex-post average. For the SVCJ models the model-based expectations match the ex-post observed variance risk premium much more closely. Again, for the 3/2 diffusion the variance risk premium is lower compared to other models of the same model class. That is, in terms of producing reasonable

Table 10: Variance Risk Premium - Full Sample

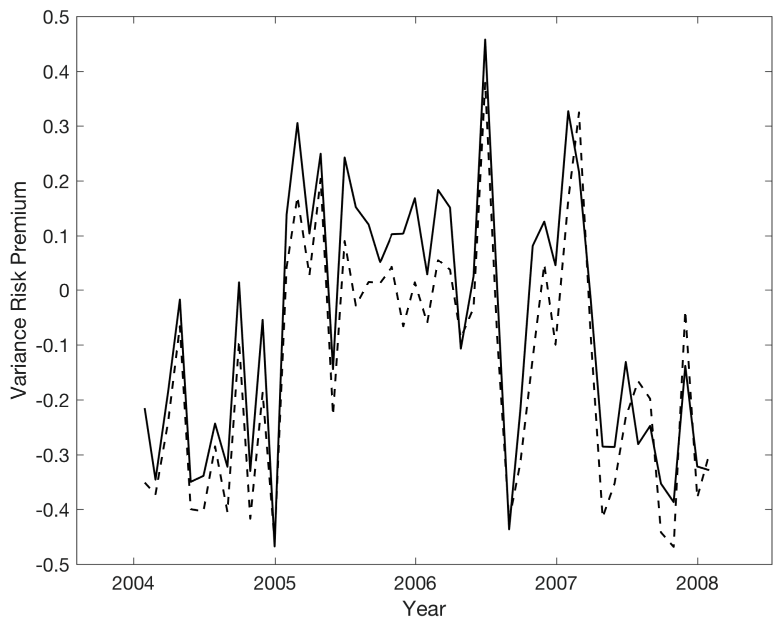
This tables provides estimation for the expected log return of the variance swap under different model assumptions. The last line shows the ex-post realization of the log return. We use monthly data from January 2000 to December 2013.

	SV Class		**SVJ Class**		**SVCJ Class**	
Models	Est.	S. E.	Est.	S. E.	Est.	S. E.
Affine Sqr	-0.338	0.043	-0.349	0.041	-0.530	0.044
Poly Sqr	-0.376	0.042	-0.374	0.042	-0.487	0.041
Affine One	-0.333	0.043	-0.335	0.043	-0.452	0.043
Poly One	-0.337	0.041	-0.344	0.041	-0.451	0.042
Affine 3/2	-0.427	0.038	-0.428	0.038	-0.638	0.043
Poly 3/2	-0.437	0.038	-0.439	0.038	-0.660	0.042
Affine Cev	-0.327	0.042	-0.340	0.043	-0.535	0.044
Poly Cev	-0.341	0.041	-0.356	0.041	-0.508	0.043
Affine Poly	-0.342	0.042	-0.345	0.042	-0.482	0.042
Poly Poly	-0.351	0.041	-0.365	0.041	-0.487	0.043
Ex Post	-0.463	0.062	-0.463	0.062	-0.463	0.062

values for the variance risk premium, SVCJ models outperform SV and SVJ models; this confirms our results in Section 4.2. Finally, our findings confirm that the diffusion specification has a much larger impact on model performance than the drift. Table 10 shows that the change in expected variance risk premia is much larger for alternative diffusion specification whereas drift specification have a minor impact.

Figure 5 contrasts the time series of the expected variance risk premia for SV and SVCJ models. For ease of exposition we restrict the time period from 2004 to 2008. We find that the risk premium expectation in SVCJ models is almost always lower than for SV models. This highlights that our results in Table 10 are not due to outliers, but that the differences are consistent across time.

SV AffineSqr vs. SVCJ AffineSqr



SV AffinePoly vs. SVCJ AffinePoly

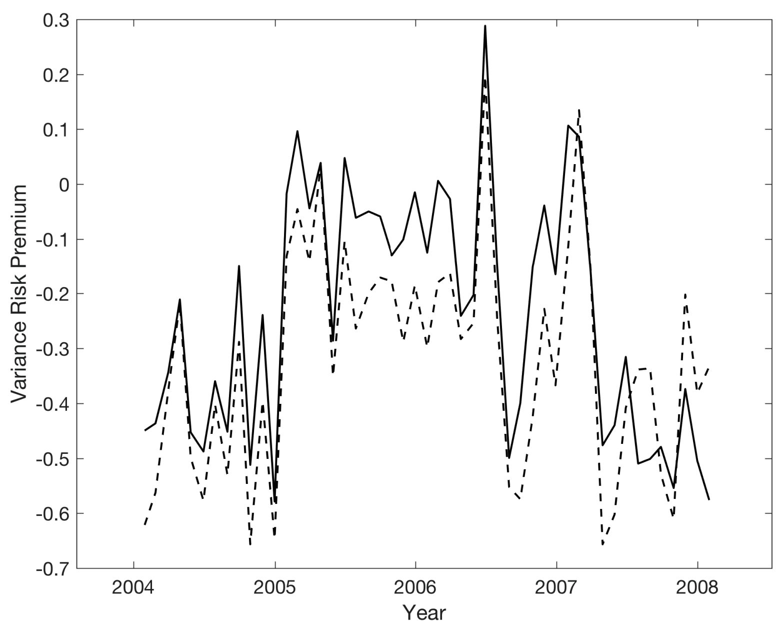


Figure 5: VRP vs. Realized Cash Flow - Lehman Crisis

This figure shows the expected log return of the variance swap for different SV and SVCJ model specifications. The time period is restricted from 2004 to 2008 to increase visibility of the differences. The top panel shows the SV AffineSqr model as solid line and the SVCJ AffineSqr model as dotted line. The bottom panel shows the AffinePoly specifications for the SV model class as solid line and for the SVCJ model class as dotted line.

5 Realized Variance as Estimation Input

In this section we analyze whether previously reported results are robust to using realized variance as an additional input in the estimation. In particular, we study whether our model ranking in Section 4.2 is robust to including high-frequency intradaily variance measures in the MCMC algorithm.[25] Similar approaches have been proposed in Jones (2003b) who uses the VIX index or Eraker (2004) who uses option prices in addition to daily return data. More recently, Maneesoonthorn, Forbes, and Martin (2017) use bipower variation as additional data in MCMC algorithms to estimate index return models. Our previous analysis resorts to realized variance only to test alternative specifications out-of-sample. Since we now use realized variance in the estimation step we no longer base our model ranking on the tests discussed in Sections 4.3 and 4.4 but focus on the DIC model ranking.

To incorporate realized variance into the MCMC algorithm described in Section 3, similar to Maneesoonthorn et al. (2017), we assume a functional relationship between model-based and realized variance:

$$\log(RV_t) = c_0 + c_1 \log(V_{t-1}) + \sigma\varepsilon_t \tag{17}$$

where ε_t is standard normally distributed. This assumption implies that RV_t is a noisy signal of the un-observable variance process which results in an additional term in the posterior distribution. Note that this specification is an approximation since for the majority of models used in our study the exact distributional relationship between the realized variance and the model implied variance is not known in closed form. Since Equation (17) is not exact, we are careful to ensure that the imposed log-linear relationship does not bias our results. To do so, we also estimate non-linear extensions such as $\log(RV_t) = c_0 + c_1 \log(V_{t-1}) + c_2 \frac{1}{\log(V_{t-1})} +$

[25] We would like to thank an anonymous referee for this suggestion.

$c_3 \log(V_{t-1})^2 + \sigma\varepsilon_t$, which are motivated by our model setup in Equation (3), as well as functional relationships based on levels. In untabulated results we find, however, that the functional form has no major impact on the overall conclusions presented in this section. We omit detailed results for these extensions to economize on space.[26] Since data for realized variance estimators are available from 2000 to 2013 and to ensure comparability to our results in Section 3.3 we use the return time series from 1983 to 2013 and add realized variances from 2000 to 2013 for model estimation.

Tables 11 to 13 provide our estimation results. We find that estimated model parameters only change marginally compared to estimates in Tables 2 to 4. As an example, the value for the leverage parameter ρ in Table 11 for the AffineSqr model is -0.6262 and in Table 2 it is -0.5997. The value of β_3 for the SVCJ AffineCev model is 1.2333 in Table 4 versus 1.1188 in Table 13. More importantly, in Table 14 we compare more formally the impact on our model rankings using the DIC criterion. We find that the model rankings change when compared to the results in Table 6. For example, the for the SVCJ model class the Affine3/2 and AffineCev swap positions in the ranking, the SVCJ AfineSqr model does outperform all SVJ models, or PolySqr is no longer the best specification for the SV model class. Our main conclusions, however, are confirmed by these additional estimation results: jump models outperform pure stochastic volatility models, non-affine specifications provide superior model fit, and Box-Cox models are outperformed by models that include non-linearities directly into the variance process.

[26]The additional term mainly affects draws of the daily variances, as we now draw V_{t-1} conditional on the return and the additional information provided by RV_t. The residual term can be interpreted as the sum of measurement and model error.

Table 11: Parameter Estimators for the SV Model Class with RV

This table shows posterior means and standard deviations (in brackets) of model parameters for drift and diffusion specifications of the SV class. The three best performing models in terms of the DIC measure as reported in Table 5 as well as the affine benchmark are presented. Parameter estimation is based on daily S&P 500 percentage returns from January 1983 until December 2013 and used the realized variance measure as additional information.

Par	**AffineSqr**	**PolySqr**	**AffineOne**	**PolyOne**	**BoxCox**
μ	0.0257	0.0280	0.0347	0.0346	0.0393
	(0.0088)	(0.0086)	(0.0086)	(0.0086)	(0.0089)
α_0	0.0241	0.0025	0.0158	0.0124	-0.0082
	(0.0023)	(0.0022)	(0.0018)	(0.0026)	(0.0023)
α_1	–	0.0054	–	0.0005	–
	–	(0.0008)	–	(0.0004)	–
α_2	-0.0208	-0.0104	-0.0106	-0.0057	-0.0204
	(0.0023)	(0.0023)	(0.0033)	(0.0034)	(0.0030)
α_3	–	-0.0002	–	-0.0012	–
	–	(0.0002)	–	(0.0007)	–
β_0	–	–	–	–	0.1766
	–	–	–	–	(0.0119)
β_2	0.1679	0.1780	0.1961	0.1959	–
	(0.0067)	(0.0070)	(0.0082)	(0.0083)	–
ρ	-0.6262	-0.6185	-0.6336	-0.6342	-0.5865
	(0.0270)	(0.0278)	(0.0275)	(0.0275)	(0.0362)
δ	–	–	–	–	-0.2409
	–	–	–	–	(0.0529)

Table 12: Parameter Estimators for the SVJ Model Class with RV

This table shows posterior means and standard deviations (in brackets) of model parameters for all drift and diffusion specifications of the SVJ class. The three best performing models in terms of the DIC measure as reported in Table 5 as well as the affine benchmark are presented. Parameter estimation is based on daily S&P 500 percentage returns from January 1983 until December 2013 and used the realized variance measure as additional information.

Par	**AffineSqr**	**PolySqr**	**AffineOne**	**PolyOne**	**BoxCox**
μ	0.0278	0.0294	0.0376	0.0379	0.0472
	(0.0087)	(0.0085)	(0.0086)	(0.0086)	(0.0093)
α_0	0.0212	0.0022	0.0139	0.0107	-0.0069
	(0.0022)	(0.0019)	(0.0017)	(0.0024)	(0.0020)
α_1	–	0.0049	–	0.0005	–
	–	(0.0007)	–	(0.0004)	–
α_2	-0.0186	-0.0091	-0.0089	-0.0046	-0.0152
	(0.0022)	(0.0021)	(0.0031)	(0.0030)	(0.0025)
α_3	–	-0.0002	–	-0.0011	–
	–	(0.0002)	–	(0.0006)	–
β_0	–	–	–	–	0.1589
	–	–	–	–	(0.0110)
β_2	0.1567	0.1676	0.1879	0.1877	–
	(0.0065)	(0.0069)	(0.0080)	(0.0080)	–
ρ	-0.6683	-0.6631	-0.6759	-0.6760	-0.6749
	(0.0261)	(0.0264)	(0.0261)	(0.0259)	(0.0335)
λ	0.0034	0.0035	0.0066	0.0067	0.0169
	(0.0013)	(0.0013)	(0.0032)	(0.0032)	(0.0074)
μ_Y	-3.1638	-3.1158	-1.5943	-1.5926	-1.0254
	(1.0531)	(1.0113)	(0.6991)	(0.7150)	(0.3873)
σ_Y	3.1617	3.1105	1.9344	1.9204	1.5889
	(0.6573)	(0.6106)	(0.3395)	(0.3257)	(0.2074)
δ	–	–	–	–	-0.2812
	–	–	–	–	(0.0535)

Table 13: Parameter Estimators for the SVCJ Model Class with RV

This table shows posterior means and standard deviations (in brackets) of model parameters for all drift and diffusion specifications of the SVCJ class. The three best performing models in terms of the DIC measure as reported in Table 5 as well as the affine benchmark are presented. Parameter estimation is based on daily S&P 500 percentage returns from January 1983 to December 2013 and used the realized variance measure as additional information.

Par	AffineSqr	Affine3/2	Poly3/2	AffineCev	BoxCox
μ	0.0429	0.0574	0.0580	0.0455	0.0467
	(0.0088)	(0.0090)	(0.0091)	(0.0089)	(0.0091)
α_0	0.0190	0.0031	0.0011	0.0081	-0.0111
	(0.0021)	(0.0010)	(0.0009)	(0.0018)	(0.0024)
α_1	–	–	0.0004	–	–
	–	–	(0.0002)	–	–
α_2	-0.0343	-0.0098	-0.0105	-0.0163	-0.0153
	(0.0031)	(0.0041)	(0.0037)	(0.0040)	(0.0025)
α_3	–	–	-0.0005	–	–
	–	–	(0.0005)	–	–
β_0	–	–	–	–	0.1412
	–	–	–	–	(0.0114)
β_2	0.1186	0.1339	0.1264	0.1619	–
	(0.0071)	(0.0065)	(0.0066)	(0.0084)	–
β_3	–	–	–	1.1188	–
	–	–	–	(0.0552)	–
ρ	-0.7137	-0.7288	-0.7710	-0.7639	-0.7034
	(0.0347)	(0.0387)	(0.0348)	(0.0353)	(0.0385)
λ	0.0137	0.0228	0.0283	0.0236	0.0099
	(0.0032)	(0.0050)	(0.0058)	(0.0060)	(0.0040)
μ_Y	-0.5984	-0.7685	-0.2448	0.1699	-1.2976
	(0.4973)	(0.3957)	(0.3416)	(0.3949)	(0.7044)
σ_Y	1.7086	1.4517	1.3530	1.4236	1.6876
	(0.2186)	(0.1520)	(0.1427)	(0.1633)	(0.2333)
ρ_J	-1.0824	-0.5510	-1.4171	-2.0430	-1.0119
	(0.2855)	(0.5167)	(0.5755)	(0.6618)	(0.9546)
μ_V	1.4482	0.6074	0.5037	0.5798	0.6078
	(0.2729)	(0.0999)	(0.0763)	(0.1157)	(0.1210)
δ	–	–	–	–	-0.1757
	–	–	–	–	(0.0701)

Table 14: Rankings of Models by DIC with RV

The table shows the Deviance Information Criterion (DIC) rankings for all models based on S&P 500 data for the time period January 1983 until December 2013. The DIC column gives the overall DIC value where lower values indicate a better model performance. The Deviance Information Criterion consists of two parts p_D the penalty term measuring model complexity and $\bar{\mathrm{D}}$ measuring model fit (see Spiegelhalter et al., 2002).

Model	**DIC**	**pD**	$\bar{\mathbf{D}}$
SVCJ AffineCev	16624.0	4347.5	12276.5
SVCJ Poly3/2	16664.2	4573.4	12090.8
SVCJ Affine3/2	17377.4	4097.0	13280.3
SVCJ AffineSqr	17556.7	3804.0	13752.7
SVJ AffineOne	18099.4	3360.1	14739.4
SVJ PolyOne	18104.3	3363.6	14740.7
SVCJ BoxCoxAffineOU	18212.0	3964.5	14247.6
SVJ AffineSqr	18236.7	3255.6	14981.1
SVJ PolySqr	18266.7	3207.7	15059.0
SVJ BoxCoxAffineOU	18570.6	3774.9	14795.7
SV PolyOne	18659.7	2904.2	15755.4
SV AffineOne	18659.8	2897.7	15762.2
SV AffineSqr	18778.1	2820.1	15958.0
SV PolySqr	18815.7	2757.2	16058.4
SV BoxCoxAffineOU	19522.7	2833.5	16689.2

6 Conclusion

We analyze the model performance of a large set of drift and diffusion specifications for modeling S&P 500 index returns. Models are ranked using complexity-adjusted return fit statistics and by comparing model-based variance paths with non-parametric high-frequency bipower variation estimates of variance. The best return fit is generated by a non-affine SVCJ model with a Cev diffusion parameter greater than one. Intuitively, such a process facilitates fast moving variances during periods of market uncertainty, and this feature leads to a superior return fit. This result is robust for various sub-sample periods.

The analysis of model-based variance paths further highlights the finding that model performance is nearly exclusively driven by the choice of the diffusion component of the SV process. Therefore, model complexity should focus on extending the diffusion function in a non-affine way with a Cev parameter of at least one. Our results indicate that sophisticated drift specifications add surprisingly little additional performance gain. Simple linear drift specifications provide sufficient flexibility and also have fewer model parameters, which may improve the stability during estimation. Further, we observe that all models perform equally well in calm market regimes, but they show considerable differences in performance during times of market stress. The comparison of realized and model-based variance confirms these findings and further supports jumps in both prices and variance. We show that jump models give more reliable estimators for the expected log return of a variance swap contract than pure stochastic volatility models. Our conclusions are robust to adding a realized variance estimator in the estimation procedure.

A Simulation Study

Table 15: Simulation Study

This table reports the parameter estimation results from a Monte Carlo study where 100 sample paths with 4000 daily returns are simulated from the true model with parameters shown as *Sim*. The simulation is performed using an Euler discretization with 100 time steps per day. The average estimated parameter of these simulated paths are reported in line *Est*. RMSE and standard errors (*StdErr*) are also reported. We use 200'000 MCMC draws with a burn-in period of 50'000 draws in every estimation run.

Para	μ	α_0	α_1	α_2	α_3	β_0	β_1	β_2	β_3	ρ	λ	μ_y	σ_y	ρ_J	μ_V
							Panel A: SV PolyPoly Model								
Sim	0.038	0.009	0.001	-0.004	-0.001	0.002	0.142	0.045	1.363	-0.614	–	–	–	–	–
Est	0.035	0.009	0.002	-0.005	-0.002	0.003	0.127	0.056	1.061	-0.601	–	–	–	–	–
RMSE	0.003	0.001	0.001	0.001	0.000	0.001	0.017	0.013	0.302	0.013	–	–	–	–	–
StdErr	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.000	–	–	–	–	–
							Panel B: SVJ PolyPoly Model								
Sim	0.046	0.006	0.001	-0.002	-0.001	0.001	0.136	0.033	1.517	-0.674	0.012	-1.231	1.681	–	–
Est	0.046	0.007	0.001	-0.003	-0.001	0.001	0.112	0.048	1.079	-0.662	0.013	-1.433	1.700	–	–
RMSE	0.000	0.001	0.000	0.000	0.000	0.000	0.025	0.016	0.439	0.012	0.001	0.208	0.024	–	–
StdErr	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.000	0.000	0.005	0.001	–	–
							Panel A: SVCJ PolyPoly Model								
Sim	0.043	0.003	0.001	-0.006	-0.001	0.001	0.096	0.059	1.185	-0.743	0.012	-0.456	1.669	-2.137	0.650
Est	0.036	0.003	0.001	-0.003	-0.001	0.001	0.099	0.048	1.021	-0.731	0.010	-0.700	1.744	-2.295	0.810
RMSE	0.007	0.000	0.000	0.003	0.000	0.000	0.006	0.012	0.166	0.012	0.002	0.292	0.080	0.245	0.161
StdErr	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.002	0.000	0.000	0.016	0.003	0.019	0.002

References

Ait-Sahalia, Y. (1996), "Testing continuous-time models of the spot interest rate," *Review of Financial Studies*, 9, 385–426.

Aït-Sahalia, Y. and Kimmel, R. (2007), "Maximum likelihood estimation of stochastic volatility models," *Journal of Financial Economics*, 83, 413–452.

Andersen, T., Benzoni, L., and Lund, J. (2002), "An Empirical Investigation of Continuous-Time Equity Return Models," *The Journal of Finance*, 57, 1239–1284.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001a), "The distribution of realized stock return volatility," *Journal of Financial Economics*, 61, 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2001b), "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96, 42–55.

Andersen, T. G., Fusari, N., and Todorov, V. (2015), "The risk premia embedded in index options," *Journal of Financial Economics*, 117, 558–584.

Bakshi, G., Ju, N., and Ou-Yang, H. (2006), "Estimation of Continuous-Time Models with an Application to Equity Volatility Dynamics," *Journal of Financial Economics*, 82, 227–249.

Bakshi, G. S., Cao, C., and Chen, Z. (1997), "Empirical Performance of Alternative Option Pricing Models," *The Journal of Finance*, 52, 2003–2049.

Bandi, F. M. and Reno, R. (2016), "Price and volatility co-jumps," *Journal of Financial Economics*, 119, 107–146.

Barndorff-Nielsen, O. E. and Shephard, N. (2004), "Power and Bipower Variation with Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2, 1–37.

Bates, D. S. (1996), "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options," *Review of Financial Studies*, 9, 69–107.

— (2000), "Post-'87 Crash Fears in the S&P 500 Futures Option Market," *Journal of Econometrics*, 94, 181–238.

— (2012), "U.S. stock market crash risk, 1926-2010," *Journal of Financial Economics*, 105, 229–259.

— (2016), "How crashes develop: intradaily volatility and crash evolution," *Working Paper*.

Benzoni, L. (2002), "Pricing Options under Stochastic Volatility: An Empirical Investigation," *Working Paper*.

Berg, A., Meyer, R., and Yu, J. (2004), "Deviance Information Criterion for Comparing Stochastic Volatility Models," *Journal of Business & Economic Statistics*, 22, 107–120.

Bondarenko, O. (2014), "Variance trading and market price of variance risk," *Journal of Econometrics*, 180, 81–97.

Britten-jones, M. and Neuberger, A. (2000), "Option Prices, Implied Price Processes, and Stochastic Volatility," *Journal of Finance*, 55, 839–866.

Broadie, M., Chernov, M., and Johannes, M. (2007), "Model specification and risk premia: Evidence from futures options," *The Journal of Finance*, 62, 1453–1490.

Carr, P. and Wu, L. (2007), "Variance Risk Premiums," *Review of Financial Studies*, 22, 1311–1341.

— (2009), "Variance risk premiums," *Review of Financial Studies*, 22, 1311–1341.

Casella, G. and George, E. I. (1992), "Explaining the Gibbs Sampler," *The American Statistician*, 46, 167–174.

Chan, K. C., Karolyi, G. A., Longstaff, F. A., and Sanders, A. B. (1992), "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *The Journal of Finance*, 47, 1209–1227.

Chernov, M., Gallant, A. R., Ghysels, E., and Tauchen, G. (2003), "Alternative models for stock price dynamics," *Journal of Econometrics*, 116, 225–257.

Chib, S. and Greenberg, E. (1995), "Understanding the Metropolis-Hastings Algorithm," *The American Statistician*, 49, 327–335.

Chourdakis, K. and Dotsis, G. (2011), "Maximum Likelihood Estimation of Non-Affine Volatility Processes," *Journal of Empirical Finance*, 18, 533–545.

Christoffersen, P., Heston, S., and Jacobs, K. (2009), "The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work So Well," *Management Science*, 55, 1914–1932.

Christoffersen, P., Jacobs, K., and Mimouni, K. (2010), "Volatility Dynamics for the S&P500: Evidence from Realized Volatility, Daily Returns, and Option Prices," *Review of Financial Studies*, 23, 3141–3189.

Conley, T. G., Hansen, L. P., Luttmer, E. G. J., and Scheinkman, J. A. (1997), "Short-Term Interest Rates as Subordinated Diffusions," *Review of Financial Studies*, 10, 525–577.

Coval, J. D. and Shumway, T. (2001), "Expected Option Returns," *Journal of Finance*, 56, 983–1009.

Drechsler, I. and Yaron, A. (2011), "What's Vol Got to Do with It," *Review of Financial Studies*, 24, 1–45.

Duan, J. C. and Yeh, C. Y. (2010), "Jump and volatility risk premiums implied by VIX," *Journal of Economic Dynamics and Control*, 34, 2232–2244.

Duffie, D., Pan, J., and Singleton, K. (2000), "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," *Econometrica*, 68, 1343–1376.

Durham, G. (2013), "Risk-Neutral Modelling with Affine and Non-Affine Models," *Journal of Financial Econometrics*, 11, 650–681.

Durham, G. and Park, Y. (2013), "Beyond Stochastic Volatility and Jumps in Returns and Volatility," *Journal of Business & Economic Statistics*, 31, 107–121.

Egloff, D., Leippold, M., and Wu, L. (2010), "The Term Structure of Variance Swap Rates and Optimal Variance Swap Investments," *Journal of Financial and Quantitative Analysis*, 45, 1279–1310.

Eraker, B. (2004), "Do stock prices and volatility jump? Reconciling evidence from spot and option prices," *The Journal of Finance*, 59, 1–37.

Eraker, B., Johannes, M., and Polson, N. (2003), "The Impact of Jumps in Volatility and Returns," *The Journal of Finance*, 58, 1269–1300.

Ferriani, F. and Pastorello, S. (2012), "Estimating and Testing Non-Affine Option Pricing Models with a Large Unbalanced Panel of Options," *The Econometrics Journal*, 15, 171–203.

Gilks, W. R., Best, N. G., and Tan, K. K. C. (1995), "Adaptive Rejection Metropolis Sampling within Gibbs Sampling," *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 44, 455–472.

Hansen, P. R. and Lunde, A. (2006), "Consistent ranking of volatility models," *Journal of Econometrics*, 131, 97–121.

Heston, S. L. (1993), "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies*, 6, 327–343.

Ignatieva, K., Rodrigues, P., and Seeger, N. (2015), "Empirical Analysis of Affine Versus Nonaffine Variance Specifications in Jump-Diffusion Models for Equity Indices," *Journal of Business & Economic Statistics*, 33, 68–75.

Jacquier, E., Polson, N. G., and Rossi, P. E. (1994), "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business & Economic Statistics*, 12, 371–389.

— (2004), "Bayesian analysis of stochastic volatility models with fat-tails and correlated errors," *Journal of Econometrics*, 122, 185–212.

Johannes, M. and Polson, N. (2009), "MCMC methods for Continuous-Time Financial Econometrics," in *Handbook of Financial Econometrics, Vol 2*, ed. Ait-Sahalia, Y. L. H., North Holland, pp. 1–72.

Johannes, M. S., Polson, N. G., and Stroud, J. R. (2009), "Optimal Filtering of Jump Diffusions: Extracting Latent States from Asset Prices," *Review of Financial Studies*, 22, 2759–2799.

Jones, C. S. (2003a), "Nonlinear Mean Reversion in the Short-Term Interest Rate," *Review of Financial Studies*, 16, 793–843.

— (2003b), "The dynamics of stochastic volatility: evidence from underlying and options markets," *Journal of Econometrics*, 116, 181–224.

Kaeck, A. and Alexander, C. (2012), "Volatility dynamics for the S&P 500: Further evidence from non-affine, multi-factor jump diffusions," *Journal of Banking and Finance*, 36, 3110–3121.

— (2013), "Stochastic volatility jump-diffusions for european equity index dynamics," *European Financial Management*, 19, 470–496.

Li, H., Wells, M. T., and Yu, C. L. (2008), "A Bayesian Analysis of Return Dynamics with Lévy Jumps," *Review of Financial Studies*, 21, 2345–2378.

Maneesoonthorn, W., Forbes, S., and Martin, G. M. (2017), "Inference on Self-Exciting Jumps in Prices and Volatility Using High-Frequency Measures," *Journal of Applied Econometrics*, 32, 504–532.

Mijatovic, A. and Schneider, P. (2014), "Empirical Asset Pricing with Nonlinear Risk Premia," *Journal of Financial Econometrics*, 12, 479–506.

Neuberger, A. (2012), "Realized skewness," *Review of Financial Studies*, 25, 3423–3455.

Ornthanalai, C. (2014), "Levy jump risk: Evidence from options and returns," *Journal of Financial Economics*, 112, 69–90.

Pan, J. (2002), "The jump-risk premia implicit in options: evidence from an integrated time-series study," *Journal of Financial Economics*, 63, 3–50.

Santa-Clara, P. and Yan, S. (2010), "Crashes, Volatility, and the Equity Premium: Lessons from S&P 500 Options," *Review of Economics and Statistics*, 92, 435–451.

Shephard, N. and Sheppard, K. (2010), "Realising the Future: Forecasting with High-frequency-based Volatility (Heavy) Models," *Journal of Applied Econometrics*, 25, 197–231.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. (2002), "Bayesian measures of model complexity and fit," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64, 583–639.

Stroud, J. R. and Johannes, M. S. (2014), "Bayesian Modeling and Forecasting of 24-Hour High-Frequency Volatility," *Journal of the American Statistical Association*, 109, 1368–1384.

Yu, J. (2005), "On leverage in a stochastic volatility model," *Journal of Econometrics*, 127, 165–178.

Yu, J., Yang, Z., and Zhang, X. (2006), "A class of nonlinear stochastic volatility models and its implications for pricing currency options," *Computational Statistics & Data Analysis*, 51, 2218–2231.

Zhang, X. and King, M. L. (2008), "Box-Cox stochastic volatility models with heavy-tails and correlated errors," *Journal of Empirical Finance*, 15, 549–566.